DOCUMENT RESUME

ED 262 976 SE 046 188

AUTHOR	Grey, D. R., Ed.; And Others
TITLE	Proceedings of the International Conference on Teaching Statistics (1st, University of Sheffield, England, August 9-13, 1982). Volume I.
REPORT NO	ISBN-0-946554-01-3
PUB DATE	83
NOTE	424p.; Edited and published by the Organising Committee of the First International Conference on Teaching Statistics. For Volume II, see SE 046 189.
AVAILABLE FROM	Teaching Statistics Trust, c/o Centre for Statistical Education, 25 Broomgrove Road, Sheffield, United Kingdom, S10 2NA, (10 pounds 50 pence).
PUB TYPE	Collected Works - Conference Proceedings (021)
EDRS PRICE	MF01 Plus Postage. PC Not Available from EDRS.
DESCRIPTORS	Calculators; Case Studies; Elementary School Mathematics; Elementary Secondary Education; Foreign Countries; Higher Education; Learning Activities; Material Development; *Mathematics Education; *Mathematics Instruction; Microcomputers; Secondary School Mathematics; *Statistical Analysis; *Statistics; *Teacher Education; *Teaching Methods

ABSTRACT

The first volume of these conference proceedings includes an introductory talk (on why statistics should be taught), three plenary papers (discussing stochastics at the school level in the age of the computer, optimum balance between statistical theory and applications in teaching, and the relevance of statistical training), and 35 invited papers. The papers are organized under nine topic headings. These topic areas are: (1) teaching statistics in schools (6-11 years age group); (2) teaching statistics in schools (12-15 years age group); (3) teaching statistics in schools (16-18 years age group); (4) teaching statistics to non-statisticians; (5) development of teaching materials at the school level; (6) training of teachers in statistics; (7) use of calculators and computers in teaching; (8) teaching statistics with the help of case studies; and (9) training statistical practitioners. Topics addressed in the individual papers summarize the situation in statistical education around the world and point the way to future developments. (JN)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED262976

U.S. DEPARTMENT OF EDUCATION
NATIONAL INSTITUTE OF EDUCATION
EDUCATIONAL RESOURCES INFORMATION
CENTER (ERIC)
☑ This document has been reproduced as received from the person or organization originating it.
☐ Minor changes have been made to improve reproduction quality.
- Points of view or opinions stated in this document do not necessarily represent official NIE position or policy.

ICOTS

PROCEEDINGS
# of the
FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME I

SE 046 188

"PERMISSION TO REPRODUCE THIS MATERIAL IN MICROFICHE ONLY HAS BEEN GRANTED BY
V. Barnett
TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

University of Sheffield
9 – 13 August 1982

ICOTS

University of Sheffield
9 – 13 August 1982

PROCEEDINGS of the FIRST INTERNATIONAL CONFERENCE ON TEACHING STATISTICS

VOLUME I

Edited by:
Organising Committee of the First International Conference on Teaching Statistics

D R Grey, P Holmes, V Barnett, G M Constable

Published by:
Organising Committee of the First International Conference on Teaching Statistics

D R Grey, P Holmes, V Barnett, G M Constable

© Organising Committee of the First International Conference on Teaching Statistics 1983

ISBN 0 946554 00 5 (set of two volumes)

ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Printed by the University of Sheffield Printing Unit

Published by Teaching Statistics Trust

Professor Lennart Råde, Chairman of Programme Committee

Professor Vic Barnett, Chairman of Organising Committee

CONTENTS OF VOLUME I

Page No.

Preface xi

INTRODUCTORY TALK

Why Teach Statistics?
V Barnett 3

PLENARY TALKS

Stochastics at the School Level in the Age of the Computer
L Råde 19

Optimum Balance between Statistical Theory and Applications in Teaching
C Radhakrishna Rao 34

The Relevance of Statistical Training
J Gani 50

TOPIC 1: TEACHING STATISTICS IN SCHOOLS (6-11 YEARS AGE GROUP)
ORGANISER: PROFESSOR K MATUSITA

Statistics in the Curriculum for Everybody - how Young Children and how their Teachers react
Tamas Varga 71

Teaching of Statistics in Elementary Schools in Japan
T Yanagawa 81

Two Examples of Statistics Activities for Elementary School Classes
Clare Heidema 91

Using Games to teach some Probability Concepts
George W Bright, John G Harvey and Margariete M Wheeler 110

Teaching Statistics in Elementary Schools in Japan
H Midzuno 116

TOPIC 2: TEACHING STATISTICS IN SCHOOLS (12-15 YEARS AGE GROUP)
ORGANISER: PROFESSOR P L HENNEQUIN

Why Teach Descriptive Statistics in Secondary Schools?
Heinrich Winter 127

The Teaching of Statistics in French Secondary Schools
G Dumousseau 144

Teaching Statistics Concepts through Instructional Games
George W Bright, John G Harvey and Margariete M Wheeler 158

A Practical Approach to Statistics
R Kapadia 169

On the Teaching of Order Statistics
K Bognár 179

TOPIC 3: TEACHING STATISTICS IN SCHOOLS (16-18 YEARS AGE GROUP)
ORGANISER: PROFESSOR J SWIFT

Amount and Kind of Statistics to be taught to 16+ year olds
F Eicker 191

Teaching Statistics to 16-18 year olds: Dreams and Reality
T Nemetz 199

Statistics for the 16-18 year olds: a Discussion
A Hanley 208

TOPIC 4: TEACHING STATISTICS TO NON-STATISTICIANS
ORGANISER: DR W G GILCHRIST

We must preach what is practised
A S C Ehrenberg 215

Teaching Statistics to Non-Statisticians - an Application oriented Approach
G Sadasivan 219

Utility through Collaboration: a Problem oriented Approach
D Taylor and G E Skipworth 232

An Open University Service Course
John Bibby 236

The Training of Service Teachers
R W Madsen 250

TOPIC 5: DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL
ORGANISER: MR P HOLMES

Development of Teaching Materials at School Level
A R Hoffer 263

Developing Materials for the Schools Council Project on Statistical Education
P Holmes 269

EDA for 16-19's?
J Bibby 277

TOPIC 6: TRAINING OF TEACHERS IN STATISTICS
ORGANISER: PROFESSOR A P SHULTE

Preparations for College Teaching in Statistics
A K Gupta 283

Teaching Design of Experiments in Indian Universities
V Bhagwandas 287

Clinical Approach to training Statistics Teachers
J Rustagi 298

Discussion of Teacher-Training Papers
A Morley 302

TOPIC 7: USE OF CALCULATORS AND COMPUTERS IN TEACHING
ORGANISER: PROFESSOR L RÅDE

The Case for using Computers in teaching Statistics
B L Joiner 307

Computer Algorithms for the Classical Distribution Functions
H G Posten 313

Use of hand held Programmable Calculators in Statistics Courses for Nontechnical Students
D R Barr 331

Simulation of Games studied in the Seventeenth and Eighteenth Centuries
J Badrikian 336

TOPIC 8: TEACHING STATISTICS WITH THE HELP OF CASE STUDIES
ORGANISER: PROFESSOR B EPSTEIN

Case Studies for an adequate understanding and interpretation of Results by Statistical Inference
M Borovcnik 341

Is a Scientific Legend a comprehensive Case Study?
V V Shvyrkov and A C Davis 356

TOPIC 9: TRAINING STATISTICAL PRACTITIONERS
ORGANISER: PROFESSOR M BENYAKLEF

Training of Statistical Practitioners
P Osinski 375

Preparing Statisticians for Government Service - a View from the United States
Marie D Eldridge and Katherine K Wallman 389

Programmes d'Enseignement et Couts de Formation
M Benyaklef 403

Les Enseignements Statistiques a l'ENSAE
Ph Tassi et A Trognon 415

CONTENTS OF VOLUME II

Page No.

TOPIC 10: EDUCATION AND TRAINING IN STATISTICS IN DEVELOPING COUNTRIES

ORGANISER: DR M N MURTHY

Statistical Education and Training in Latin America: a General Appraisal
Enrique Cansado 431

Statistical Education and Training in the Developing Countries in Africa
J N Adichie and Biyi Afonja 439

Statistical Education and Training in Developing Countries of Asia and the Pacific
M N Murthy 450

Comments on the above Papers
K T de Graft-Johnson 483

Report on session on Statistical Education and Training in Developing Countries
B P Adhikari 487

TOPIC 11: CO-OPERATION BETWEEN ACADEMIC AND PRACTISING STATISTICIANS

ORGANISER: PROFESSOR W G HUNTER

Co-operation between University and Industry Statisticians in the United States
R D Snee 491

Some Observations from Europe
Lars-Erik Öller 503

Building Bridges between the Academic and Real Worlds - some Observations from South America
C A Peres and F A Morettin 509

Building Bridges between the Academic and Real Worlds
S Jones, B J T Morgan and G B Wetherill 521

Closing the Gap between Statisticians and Statistics Users
Guido E del Pino and Carlos Prado-Campos 527

Co-operation between Academic and Practising Statisticians
A Prat Bartes 536

TOPIC 12: TEACHING SURVEY SAMPLING
ORGANISER: DR D SINGH

The Teaching of Sample Design to Government Statisticians from Developing Countries
Joseph K Garrett 541

Teaching of Survey Sampling
D Singh 552

Teaching of Survey Sampling
P K Bose 566

TOPIC 13: TEACHING DESIGN AND ANALYSIS OF EXPERIMENTS
ORGANISER: DR D A PREECE

The Design and Analysis of Experiments - what is it about?
D A Preece 575

Teaching Experimental Design in the Computer Age
R Mead 591

Teaching Concepts of Design and Analysis using Real and Simulated Experiments
J R Elliott and J C Robinson 603

Teaching Experimental Design: prescribed by a Medical Statistician
Sheila M Gore 627

Contribution to Discussion of the above Papers
B Jones 633

TOPIC 14: TRAINING INDUSTRIAL PRACTITIONERS
ORGANISER: DR R D SNEE

The Polymath Consultant
A A Greenfield 641

Teaching Applied Statistics for and/or in Industry
H C Hamaker 655

Practicals, Case Studies, Projects and Professional Experience: some Methods of developing Skills among Budding Statisticians
G K Kanji 671

Training in Statistical Quality Control
A F Bissell 685

Discussion of Papers by Greenfield, Hamaker and Kanji
A F Bissell 691

Discussion of Papers by Greenfield, Hamaker and Kanji
D G Watts 695

TOPIC 15: PRINCIPLES OF LEARNING PROBABILITY AND STATISTICS
ORGANISER: PROFESSOR E FISCHBEIN

Learning from Experience - Bayes' Theorem: a Model for Stochastic Learning Situations?
H Steinbring and G von Harten 701

Children's Choice Behaviour in Probabilistic Situations
Ruma Falk 714

An Idea of structuring Probability Teaching based on Dienes's Six Stages
Anna Olecka 727

Does the Teaching of Probability improve Probabilistic Intuitions?
E Fischbein and A Gazit 738

Stochastics Teaching based on Common Sense
H-J Bentz 753

A Survey of Probability Concepts in 3000 Pupils aged 11-16 Years
D R Green 766

Misconceptions of Probability, systematic and otherwise; teaching Probability and Statistics so as to overcome some Misconceptions
J M Shaughnessy 784

The Role of Explanation in Teaching Elementary Probability
J L Myers 802

Probabilities: Learning by doing or learning by imitating
A Abele 813

TOPIC 19: TEACHING OF AGRICULTURAL STATISTICS
ORGANISER: DR D SINGH

Training of Personnel for Agricultural Statistics in Jordan
Z Takrouri 831

PREFACE

The first International Conference on Teaching Statistics (ICOTS) took place in the University of Sheffield over the period 9-13 August, 1982. The Conference attracted much interest from all corners of the world. Over 400 full participants from more than 60 countries (accompanied by 100 or so associate participants) spent an active week attending Plenary Sessions, Invited Talks, Contributed Paper Sessions, Workshops and a variety of social events including walks in the Peak District hills and a reception and buffet generously provided by the Lord Mayor of Sheffield on behalf of the City.

The formal programme was very full and covered a wide range of interests in statistical education and training, including the teaching of statistics for all age groups in schools and colleges, the training of practitioners in different applied disciplines, the special problems in developing countries and the difficulties of teaching specific topics. An opening address was given by Professor S. Gregory, Pro-Vice-Chancellor of the University of Sheffield and the closing session was chaired by Mr. H. Ben-Amor, the Director of the UNESCO Office of Statistics.

The Conference was sponsored by the International Statistical Institute and the organisers are grateful to the ISI, and to UNESCO, for financial support. The technical programme and local organisation were arranged by the Programme Committee and the Organising Committee under the respective chairmanships of Professor L. Råde and Professor V. Barnett. The full membership of the two committees was as follows:

Programme Committee

L. Råde (Chairman)
E. Cansado
P. Holmes
B. L. Joiner
M. N. Murthy
G. E. Noether
J. O. Oyelese
B. Penkov
R. Zielinski

Organising Committee

Professor V. Barnett (Chairman)
G. M. Constable
D. R. Grey
P. Holmes
A. Barnett (Conference Secretary)

This Conference arose from the special interests and actions of the International Statistical Institute (ISI): in particular, from the establishment in 1978 of an Education Committee Task Force charged with organising an international conference on the teaching of statistics. The field of statistical education has become a very extensive one and the need for international conferences on the teaching of statistics has been emphasised on several occasions, especially at the many round table meetings on the teaching of statistics arranged by

ISI, and at the international conferences on mathematical education organised by ICMI (International Commission on Mathematics Instruction).

These published Proceedings of the first ICOTS consist of all plenary and invited papers together with contributions from invited discussants, collected together under appropriate topic headings. It is regrettable that the sheer volume of this material has precluded the possibility of including any of the very worthy contributed papers.

It is appropriate to acknowledge the vast effort extended by all members of the Programme and Organising Committees in setting up the Conference and to thank all Topic Organisers for helping to provide an interesting programme. In addition to thanking all authors for their co-operation in submitting papers in suitable form for publication, it is necessary to express particular indebtedness to three people who have given invaluable assistance in bringing this publication into being: Dr. David Grey, who organised the overall preparation of the Proceedings, Mrs. Jenny Cameron, who has undertaken the mammoth task of re-typing all the papers in standard format, and Mr. Bill Seaton, one of the Conference participants, who during an extended period of sabbatical leave in Sheffield, has willingly devoted many hours to the onerous task of proof-reading.

In accordance with views expressed by Conference participants it is hoped that this Conference will be the first of a series of such international conferences on teaching statistics to be held at four-yearly intervals.

V. Barnett
(Chairman, Organising Committee)

Sheffield, December 1982

INTRODUCTORY TALK

WHY TEACH STATISTICS?

V Barnett, University of Sheffield

It might appear a little strange that the opening comments in the First International Conference on Teaching Statistics (ICOTS) should seem to question the very basis of that conference. Such ambiguity of theme is even more perverse from someone who has for more than twenty years been thoroughly immersed in efforts to awaken interest, or change attitudes, to the teaching of statistics at the national and international levels and across the spectrum of school, university and professional work.

For a long time I have been as vocal as anyone in stressing the need for an international conference solely and specifically on statistical education. Of course, the teaching of statistics has featured as a topic of earlier conferences - I would remind you of the invited-participation conference on pre-college statistics teaching at Carbondale, Illinois in 1969 (Råde, 1970). Then there were the round-table meetings held in conjunction with ISI sessions: in Vienna in 1973 (Råde, 1975) and Warsaw 1975 (Breny, 1976) both on school-level statistical education and concerned with general principle more than classroom implementation. A few sections of other ISI conferences have considered specific aspects of teaching or training: e.g. of government statisticians, at the New Delhi session in 1977. Detailed consideration of the teaching of mathematics takes place at the quadriennial meetings of the International Congress on Mathematical Education. Sections at Exeter (1972), Karlsruhe (1976) and Berkeley (1980) were devoted to statistical education, but with very limited coverage in respect of time, topic and level.

The springboard for ICOTS was the New Delhi ISI session of 1977 where a strong initiative developed for a truly international and wide-ranging conference on teaching statistics. The ISI Education Committee was restructured and set up task forces to take positive action on different aspects of statistical education. One of these task forces (under the Chairmanship of Professor L. Råde from Gothenburg, Sweden) was charged with the task of planning ICOTS, and proposed that the ISI should invite the Department of Probability and Statistics at Sheffield University to organise the conference; in acknowledgement of its long-term and wide-spread commitment to the study and development of statistical education. We were very honoured to receive this invitation.

And so we find ourselves together this week to share experiences, to learn new ideas and to hear of individual approaches in the context of a vast range of different national backgrounds (over 50 countries are represented here) and on a broad canvas of topics and emphases in statistical education. The specialist, and the generalist, is catered for with sessions on all levels of school and college work, development of teaching materials and use of calculators, computers, case studies and so on. We will be considering the special needs of developing countries; the teaching of statistics to non-statisticians as well as statistical practitioners, in industry and in other spheres; the training of teachers, and the teaching of specialised topics such as survey sampling and experimental design. Since I have not been concerned with drawing up the programme I feel that I can congratulate the programme committee on preparing such a veritable feast (might I say orgy) of unremitting delight and enlightenment. I hope it does not lead to indigestion!

This conference is an exciting prospect! I hope it may be only the beginning of a series of international meetings which attempt to bring together the wide assortment of valuable experiences of different individuals and nations in furthering the cause of sound statistical education. Too many separate attempts have been made to 'invent the wheel' in this matter - too many of the invented wheels have been square. Let's begin to work jointly to make a round wheel that we can push along together for mutual benefit. We are approaching a time when every citizen in every country must be given some rudimentary understanding of statistical ideas simply to cope with the need for numeracy in modern society.

The conference comes at an interesting time in the development of statistics and of statistical education. In the United Kingdom we can look back on some distinct landmarks rather interestingly spaced about 25 years apart.

The Royal Statistical Society is to celebrate its 150th anniversary in 1984. About 125 years ago, an energetic and determined woman started a lone crusade for the publication of extensive social statistical data and in particular for the establishment of a professorship in Applied Statistics at Oxford University, ending with a cogent plea on this matter to the famous geneticist Francis Galton almost exactly 100 years ago. She was known at the time as the 'passionate statistician' and never saw the fruits of her plea - in fact, Oxford established its first chair of Statistics in 1948. Some of you may have heard of this lady under a different name : as 'the lady of the lamp' - Florence Nightingale, whose successes in pioneering nursing practice are better remembered than her efforts for statistical education. 1984 also sees the centenary of Karl Pearson's appointment to a Professorship in Applied Mathematics at University College London but it was a further 25 years or so before he was to become the head of the new (indeed first) University Department of (Applied) Statistics. At this same period (about 75 years ago) we were seeing in the U.K. the start of an avalanche of practical statistical applications in such vital areas as biology, agriculture and industry. It was around 50 years ago that the statistical world began to reverberate to the critical conjunction of Jerzy Neyman and E.S. Pearson. Both were engaged at University College London on their crucial work on the Neyman-Pearson theory of inference and E.S. Pearson took over the headship of the Department of Applied Statistics in 1934 (and soon also found the same hallowed corridors stalked by the famous R.A. Fisher in the role of Galton Professor of Eugenics). Even the presence of statistics at school-level is approaching its quarter-centenary and the single statistic that 80% of secondary schools in England and Wales now teach some statistics is an impressive, if somewhat arbitrary, reflection of changes in attitudes and emphasis to statistics as a subject of study.

Of course, different nations have had different experiences. In Hungary, statistics (or at least probability and combinatorics) have been in and out of the secondary school curriculum for over 130 years, whilst it would appear that Japan (in spite of the valuable contributions its nationals make to statistical theory and practice) has no separate curriculum for statistical education at school-level (some topics are taught in the mathematics lessons) and no independent departments of statistics in the Universities. I look forward to hearing in the later sessions of recent developments in Japan. Statistics does not feature even in the school mathematics courses in many countries, whilst in contrast in Sudan 'basic statistics is in the syllabus of a subject that must be passed by a student wishing to enter [any subject of study in] higher education' (Downham and Gattar, 1982). Unfortunately it is all too easy for pupil and even teacher to ignore the statistics component. My role as Chairman of TOTSAS, the ISI taskforce on teaching statistics at school level, has given me a unique opportunity to witness the enormous range of different attitudes to, and levels of incorporation of, statistics in schools in different countries. We have just produced a book entitled Teaching Statistics in Schools throughout the World (Barnett, 1982). This was published by the ISI only in the last week or so and copies are available at the conference should you wish to explore some of the widely assorted systems and practices in different countries.

But let me return to my main theme. In the light of the various positive remarks I have just made, can there really be any doubt of the value of teaching statistics or of holding a major conference on such a topic. As far as this gathering is concerned the answer must surely be a firm NO : we must teach statistics more comprehensively, more widely, more wisely, more realistically, more everything!

As a secondary-school class-room teacher of statistics you will be very concerned to extend and improve (or even just to introduce) the teaching of the subject. As a teachers-training-college lecturer in statistics you will be anxious to broaden the exposure of student teachers so that their influence can be felt in the schools. If you are concerned with statistical training in medicine, industry or government you will want to ensure that practitioners can contribute soundly on important and real problems and that their organisations are well aware of what the statistician has to offer. If you are engaged in statistical curriculum development at any level you will be anxious to improve teaching materials and syllabus content and to disseminate the results of your work to a wide audience.

Notice, however, that the emphasis in my remarks is on what you as a statistical educator will want to do, and statistics is seldom a subject which controls its own destiny at any level or in any educational or national environment. It is a very different, and very crucial, matter to establish the need for statistical instruction or for the use of statistical techniques in the wider frameworks of a school curriculum, in applied subjects at college level or in the operations of a business or government organisation. Why should we teach statistics? - is very much the automatic response to any proposal to do so, and it is a natural and proper response. There is no intrinsic reason why a school or college should offer statistics or why an individual should wish to study it.

I am sure that many of you have had experiences which are not so different from mine in this matter.

As a mathematics undergraduate my head of department (one of the most eminent pure mathematicians of his generation) repeatedly asserted that statistics was an inferior subject not worthy of the attention of anyone with claims to mathematical ability. It was this attitude which was paramount in my decision to become a statistician! Teaching at university level I have had to cope with many mechanical engineers, agriculturalists, physicians (and so on) who saw no good reason to attach importance to service courses in statistics for their students. The grudging acceptance of a 10 lecture course on the whole of statistics (including Markov chains and factor analysis) is very much a feature of applied departments at college level or even in teacher-training. And as far as the 'man on the Clapham omnibus' is concerned the lay attitude to statistics is one of suspicion

'You can prove anything by statistics'

'Lies, damn lies and statistics' etc.

This is a worrying state of affairs, since so few areas of social or personal life can afford to ignore what statistics has to offer - the very reading of a newspaper or watching of a television programme or of advertisements requires some degree of statistical understanding. And of course such an attitude of resistance or even of hostility is bound to be reflected in the way in which the educational system responds to the incorporation of statistics. Some countries are beginning to see an easing of such suspicion and a growing acceptance of the need for statistical instruction at different levels, whether this is for the production of statistical experts or merely to 'arm the citizen against the media'. But in contrast, some countries have not even reached that level of public awareness of statistics for any resistance to have yet developed. They face the double problem of trying to introduce ideas of statistics at the same time as reassuring their society of its respectability.

Such considerations imply that the question: Why teach statistics? must be a crucial one in any attempt to introduce or extend the teaching of statistics, or in any effort to encourage widespread use of statistical methods. the onus of justification must rest with the statistician; as practitioner, teacher or educationalist. And in my view that justification must be based on one clear principle - that statistics is a practical discipline for understanding the indeterministic world that we live in and for solving the real problems in society from agriculture, through meteorology to zoology - from A to Z! The message is not that statistics might be just another topic of possible study, but that it is central and essential to the education and training of the school-child, the university student, the doctor or even the linguist. The statistician has to be more evangelical in this matter, whether he is a junior teacher in a mathematics department trying to persuade sceptical colleagues to see the place of statistics in the mathematics programme, or whether he is trying to influence central governmental attitudes on the place of statistics in the school curriculum or in society generally.

Thus the main response to the question is 'because we cannot afford not to', and we have to persuade others (including some statisticians and mathematicians) to believe this. But the question has a much more fundamental role for the teacher of statistics. Whether contemplating how you as a teacher will present a particular syllabus, or as an educationalist what recommendations you should make for course design, training programmes or curriculum development, the starting point has to be the same. Before considering how to teach some prescribed syllabus or even what should be the content of a syllabus, we need the clearest possible notion of just why we are teaching statistics at this level, to this group of students etc. Only after we have faced this basic question can we hope to make sense of such matters as the choice of content, teaching method, reinforcement technique, motivation and assessment.

Take a simple example - the teaching of the binomial model and distribution. The approach will depend on many factors. Is it the first introduction of the topic to 14 year olds in a general studies class or to 18 year olds in a mathematics lesson, or is it part of the statistical methods component of a university geography, or sociology, course: is it in a statistics course for secondary school student-teachers or does it feature in a refresher course on quantitative methods for doctors or for junior managers in industry? The level and method of presentation must be carefully matched to the student group in respect of their maturity and current state of knowledge; it makes no sense to use the binomial distribution as a vehicle for teaching the mathematics of the weak law of large numbers to 14 year olds in a general studies, or even mathematics, lesson. But again, even more important is the question of why such a group, at such a level is being taught such a topic : the binomial distribution.

I can recall my experiences in the mid-sixties as an examiner on a Pure Mathematics with Statistics A-level GCE examination (the public examination of 18 year olds in the U.K.). In various years there would be a question which might typically start

"Define the binomial distribution..."

and then go on to ask about its practical application, or for the solution of a practical or numerical exercise involving the binomial distribution. It was a depressing experience to find candidate after candidate responding:

"The binomial distribution is the expansion of $(p+q)^n$..."

What has such an answer to do with probability and statistics? Obviously such an irrelevant response was reflecting how the candidates were taught the topic (not in all schools, but uncomfortably many) - with no regard for the nature of a probability distribution or for its crucial role in modelling real problems. Instead, it was being presented as a convenient tool to illustrate a mathematical expansion. Why should this attitude be taken? We can only hazard a guess, but my conclusion was that it was not due to a conscious desire to conceal the importance of the binomial distribution as a model - to deliberately manipulate its use to the service of some mathematical exercise. Sadly it was because no-one (curriculum developer, examiner or teacher) had provided a clear enough answer to why the binomial distribution was being taught in the syllabus. Too few teachers at that time could provide the answer from their own education or practical experience - they probably had to keep one step ahead of the class from some equally formal and inadequately motivated textbook. (The dual problems of teacher-training and availability of training aids are vital - I am sure they will figure widely in the discussions at this conference).

But let us stay for a moment with the example of teaching the binomial distribution and see how different purposes in teaching this topic produce quite distinct approaches to suit different needs. We will go through the groups in order.

i) '14 year-olds in a general studies class'. For this group there can be no question of any mathematical treatment. We need to awaken interest through many practical examples (winning or losing a game of tennis, answering 'yes' or 'no' to a question, passing or failing a test, 'rain' or 'fine' tomorrow) in the notion of random variation in a dichotomous situation : that in six games of tennis the number of wins is random (it might be six, or one, depending on who plays who, and on what occasion!). Statements such as '8 out of 10 dogs prefer Dognosh' or 'one in every four children is Chinese' begin to make sense and even primitive notions of estimation develop : 'If 90 people out of a sample of 102 have a radio, perhaps about 90% of the population do so'.

(This sort of awakening of statistical understanding for 11-16 year olds across the curriculum, for all abilities, and tied totally to everyday problems, was the basis of the Schools Council Project on Statistical Education which was directed over 6 years in Sheffield by Peter Holmes. The publications include 27 teachers' and pupils' units, 'Statistics in Your World', and books entitled Statistics in Schools 11-16 : a review (Holmes et al, 1981) and Teaching Statistics 11-16).

ii) '18 year olds in a mathematical lesson' can cope with a more mathematical (algebraic) treatment - with the notion of a distribution and its mean, and an estimator. But again it is crucial that the empirical basis is stressed and illustrated and that the model can be applied to the study of real problems (whether cars on ferries, or bacteria in test tubes). I have no sympathy at this level with probability or statistics being exploited merely to illustrate mathematical concepts.

The most grotesque example I recall of such use was a GCE A-level question along the lines:

A continuous random variable X has probability density function

$$f(x) = a + bx + cx^2 \quad (0<x<2).$$

The mean and standard deviation of X are 8/5 and 1/5, respectively.

Determine a, b and c.

This is purely an exercise in solving three linear equations, from which we obtain $a = 3/20$, $b = 3/20$, $c = 15/40$. Who would want to use such a probability model (especially with $f(x)$ negative for $x < 0.46$).

This is not to deny the importance of mathematics in statistics - it is vital. I merely question the teaching of statistics as if it were mathematics.

But let us move on briefly through the other groups:

In the

iii) 'statistical methods component of a university geography course', the binomial distribution has to be seen to be important to the geographer. Examples for motivation and analysis must be specific.

There are some teachers of statistics who hold the view that it should not be necessary to choose practical examples which are closely related to the interests of the students - they would claim that statistics is so clearly important that whether we are teaching an economics student or an engineering student the same example on the distribution of earthworms will serve equally in both cases. I totally disagree with this view; my experience convinces me that fully specific practical illustrations are vital and must be carefully chosen to add excitement and relevance to the teaching. We can illustrate this for those other groups for whom we contemplated using the binomial model.

In the case of

iv) 'statistical methods for university sociology students' the typical problem is to generate some interest in what students may feel is a rather tedious short extra course. A useful vehicle could be the idea of randomised response or the 'principle of the irrelevant question'. For example, we might describe a survey on some sensitive issue - the use of a drug like marijuana or 'pot'. To ask respondents directly 'do you smoke pot?' invites a dishonest or indignant reply. Instead we build up confidence by playing a game with a randomising device e.g. a box with red and green balls, where just one appears in a window at random. We then ask the respondent to answer a questionnaire such as the following

---

SHAKE THE BOX AND OBSERVE THE COLOUR OF THE BALL

IF RED, ANSWER QUESTION 1; IF GREEN, ANSWER QUESTION 2

REPLY YES OR NO.

Q1 IS YOUR BIRTHDAY IN APRIL?

Q2 DO YOU SMOKE POT?

YES/NO

Apart from the practical advantages of instilling confidence through confidentiality and avoiding legal problems of having specific information about the individual, this provides a most compelling use of the statistics topic; the binomial model. If we know the proportion of red and green balls, and of birthdays in April, we obtain an immediate binomial estimate of the proportion who smoke 'pot'!

Peter Holmes, on his Schools Council Project for 11-16 year olds, tells me that he used this approach most successfully with a class of 12 year olds on the issue of smoking cigarettes during the school lunch break. A quick estimate (incorporating a bit of luck) of precisely 4 smokers led to 4 red faces and a class clamouring to hear more and more about statistics!

For the

v) 'secondary school student-teachers' the challenge is a double one of having the experience to set the topic in any of a range of practical situations (depending on need) and the skill to teach the topic in a compelling way within that situation. All the teacher needs is a degree in statistics, 3 or 4 years practical experience in two or three different fields of application, a postgraduate qualification in education and a charismatic personality! Unfortunately, all the teacher often possesses is uncertainty bred of a poorly presented 10-20 lecture introductory course at College and an inadequate text book!

For the

vi) 'doctors' or 'junior managers' the task is much easier in principle - although skill is still required on the part of the instructor. The audience is already self-motivated. As practitioners they will respond to anything that increases their power to operate. So the binomial distribution needs merely to be offered in this spirit : e.g. as the basis of clinical trials analysis for the doctor or of sampling inspection techniques for the industrialist. This works, as long as you convince the audience that you really do know what you are talking about in practical down-to-earth terms.

So how am I to answer the question : why teach statistics? Essentially in the following way!

We are surrounded with problems - social, economic, scientific, personal. Almost without exception they involve indeterminism, the element of chance. To sovle such problems we need to understand probability and statistical method. We teach statistics to give the individual (depending on need and ability) a level of understanding or expertise in this matter of problem-solving. It might be thought a simple answer, but if whole-heartedly accepted it has profound implications for a wide range of further considerations such as - who should teach statistics, to what groups of pupils or students, for what purpose, on what syllabus, etc. To spell out these implications in detail I would need to write a book, rather than to present an introductory talk.

Consequently, I shall conclude by taking one or two specific topics or themes germane to the different sessions of the conference as a basis for some purely illustrative comments.

A. Teaching Statistics in Schools (12-15)

Society is placing ever growing demands on citizens in respect of their numeracy : their ability to assimilate and interpret quantitative information. Such information is becoming increasingly statistical in nature and statistical edcuation for this age range must face up to this need. We should not be aiming to produce precocious statistical experts but should be seeking to instill, across the curriculum and ability range, some basic understanding of statistical reasoning and informal inference. This is best approached through real practical problems from everyday experience. The problems are manifold. Where do we find appropriate teaching material? Who can cope with such an empirical, look-and-learn approach to teaching? How is it to be fitted into the school programme? The Schools Council Project on Statistical Education had just these problems to face in producing its units for pupils and teachers aimed at arming the future citizen to cope with the outpourings of television, newspapers, government, etc. (See Schools Council Project on Statistical Education, 1980a, 1980b; Holmes *et al*, 1981).

Some groundwork may have been laid in the primary schools (6-11 year olds) in respect of collection, tabulation and presentation of data, and the citizen-arming process will pave the way to the *variety* of needs at the next level (16-18 year olds - see below). One suggestion taken up in the recent report of the U.K. Committee of Enquiry into the Teaching of Mathematics in Schools (Cockcroft, 1982, *Mathematics counts*) is an interesting one - that there might be a 'statistics coordinator' in a school who acts as a focus for the statistics teaching across different subjects and ability levels.

B. Teaching Statistics in Schools (16-18)

Here we have a new dimension to face. There are quite different needs for distinct groups of pupils. It is common to think that the mathematically oriented and able are the important ones to cater for. But they represent a minority and are the easiest to deal with. If they have had appropriate earlier motivation, they can be given a more mathematical consolidation prior to further studies at college or university. Their syllabus is relatively straightforward - as is choice of teaching method (still with empirical reinforcement) and there is hope of finding at least some mathematics teachers who can present the material appropriately. One danger for this group is the universal overestimation of how much material can reasonably be presented. Syllabuses, which their designers claim to be of reasonable extent, can often on detailed scrutiny be shown to contain as much as 100% more implicit content than a parallel (and supposedly equal) course in some other subject e.g. physics.

But there are two other larger groups at this age which present a greater challenge and are even less well provided for. There are those in schools and colleges who are either studying purely qualitative or craft/vocational subjects who need a more sophisticated version of the 'citizen-arming' approach proposed for the 12-15 year olds, and, in contrast, those who are studying quantitative or technical subjects (including economics, business studies; even biology and geography) who need to be introduced to *relevant statistical method* without unnecessary mathematical frills and taught with copious reference to real-life data-based problems. But again, we must ask whether we have teachers capable of teaching in this way, at this level, in the different subjects and whether we have appropriate teaching material? A start on this problem in the U.K. is being made in the new Sheffield-based Leverhulme project for 16-19 year-olds studying business or technological subjects in Colleges of Further Education; or economics, biology, geography, etc. in schools. As with the earlier Schools Council Project, the Leverhulme Project (STEP 16-19) is again being directed by Peter Holmes. A detailed and lengthy pre-project report on the education and training of 16-19 year olds was prepared two to three years ago for the Centre for Statistical Education in Sheffield University (see Barnett *et al*, 1979).

C. Teaching Statistics to Non-statisticians

The question of who should teach statistics 'service courses' has always been a thorny one - should the biologist teach statistics to his own students or invite the statistician to do so? Relevant illustrative material is vital; so is sound well-constructed statistical argument. At school-level the former consideration (and time-tabling problems) may support the subject specialist; beyond school level I believe it is the statistician who can do the better job provided time is invested in accommodating the special subject realistically. The major policy committee in our university recently concluded 'that all teaching, provided it is directed to the particular needs of its recipients, is best conducted by those who are specialists in the subject'.

D. Training of teachers, development of teaching materials, use of teaching aids

Here we have absolutely central considerations which feature widely in the conference. Any approach to teaching statistics, but particularly the empirical practical style I favour, depends on quality in these three matters. Sadly, no country can feel particularly proud of its efforts in teacher-training or materials-development for statistics, and much investment is needed. There are real problems in teacher-training, not least of which is how to find *enough time* (in training or in service) to convey the right attitude and experience to potential teachers of statistics. There is also a fundamental problem : statistics is not the usual linear discipline where we teach by placing one brick of understanding on top of another already firmly placed one. Instead we must persuade teachers to adopt a 'window-gazing' approach - first showing something of the breadth of application of statistics and gradually focussing down on more and more specific aspects as students develop gerater confidence, familiarity and experience. This is no easy task.

As far as teaching-materials are concerned, we need to produce high quality well-reasoned soundly-motivated texts*and resource matter, and must not hesitate to discard much inadequate matter which is hampering progress. We must also be alert to an insidious problem : calculators and computers can be important in teaching statistics, but they are merely tools, not teacher nor text!

E. Curriculum Development

The 1960's and 1970's saw an avalanche of activity in curriculum development across all subjects. Money flowed like water (or at least crude oil) and some of it even slopped over statistics. It seems unlikely that we shall return soon to a liberal financial climate in education and yet we have much to do in the development of syllabus and courses (or, more fashionably, programmes) in statistics. I see a major difficulty here. Shortage of financial support for research and development in statistics will certainly restrict the detail in which courses and materials are produced and virtually eliminate vital stages of testing, validation, reappraisal etc.

To put this in perspective we might note that the two books and 27 teaching units of the Schools Council Project on Statistical Education took about 20 man years of effort. Depending on your stand-point it cost about:

1 000	Gold Krugerrands
4 000	Gold Sovereigns
110 000	Dinar
220 000	UK£
400 000	US$
1 000 000	DM
2 500 000	F Francs
6 600 000	Schillings
42 000 000	Pesetas
100 000 000	Yen
500 000 000	Lire

- a lot of money by any standards!

But the reduction of support funds will also encourage the 'statistics on a shoestring brigade'! In my capacity as Chairman of the ISI taskforce TOTSAS I have been involved in many discussions in different countries on how they might introduce a programme of statistics teaching in schools. On three separate occasions the attitude was identical - to call together a *small* committee representing university statisticians and the national statistical body (and perhaps an odd teacher) and to ask them to meet on 3 or 4 occasions to draw up 'the statistics syllabus'. The result is always the same (we saw it in the U.K. twenty five years ago!) A page of topic headings such as might describe *any* introductory course at *any* level, for *any* group and with any emphasis, which is repeatedly misinterpreted and misapplied until a decade later someone asks 'where did we go wrong?' The error in my view was in not asking and seriously answering the very question with which I started this talk - *just why are we wanting to teach statistics in the first place?* Let me leave you to think further on this matter. Whether or not you agree with my answer to the question, I hope that it might provide a springboard for discussion and debate in many of our sessions throughout the conference.

References

[1] Barnett, V. (Ed) (1982) *Teaching Statistics in Schools throughout the World*. International Statistical Institute; Voorburg, The Netherlands.

[2] Barnett, V., Cable, P., Conway, D.A., Harris, J., Holmes, P., Mann, E.J., Morley, S.A., Oldfield, S.J. and Wilkinson, R. (1979) *Statistical Education and Training for 16-19 year olds*. Centre for Statistical Education, University of Sheffield.

[3] Breny, H. (Ed) (1976) *The Teaching of Statistics in Schools*. International Statistical Institute; Voorburg, The Netherlands.

[4] Cockcroft, W.H. (Ed) (1982) *Mathematics Counts. Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr. W.H. Cockcroft*. HMSO, London.

[5] Downham, D.Y. and Gattar, H. (1982) Chapter 12 of Barnett (1982).

[6] Holmes, P., Kapadia, R., Rubra, G.N. and Turner, D. (Ed) (1981) *Statistics in Schools 11-16: a review* (Schools Council Working Paper 69). Methuen Educational, London.

[7] Råde, L. (Ed) (1970) *The Teaching of Probability and Statistics*. Almqvist and Wiksell, Stockholm.

[8] Råde, L. (Ed) (1975) *Statistics at the School Level*. Almqvist and Wiksell, Stockholm.

[9] Schools Council Project on Statistical Education (11-16) (1980a) *Teaching Statistics 11-16*. Foulsham (for the Schools Council), Slough, U.K.

[10] Schools Council Project on Statistical Education (11-16) (1980b) *Teachers and Pupils Notes. Statistics in Your World*. Foulsham (for the Schools Council, Slough, U.K.

PLENARY TALKS

STOCHASTICS AT THE SCHOOL LEVEL IN THE AGE OF THE COMPUTER

Lennart Råde, Chalmers University of Technology, The University of Gothenburg

INTRODUCTION

Programmable machines, from programmable calculators and microcomputers to large computer systems, are extremely important in statistical work today. This is one reason why the use of such machines should be considered in stochastics (probability and statistics) courses both at the school and the college level. Another, maybe even more important, reason is the possibility these machines offer for a concrete and experimental approach to the teaching of probability and statistics, especially at the school level. The programmable calculator is especially important in this respect because of its low cost (comparable to that of a book) and its availability. It allows each student to have access to her or his own "pocket computer".

In this paper mainly two examples will be used to demonstrate the flavour of a computer oriented approach to the teaching of stochastics at the school level.

THE TWO TOWERS, AN INTRODUCTORY EXAMPLE

Two towers are built of m and n blocks. A tower is chosen at random and a block is moved from this tower and put on top of the other tower. This procedure is continued until all the blocks have been moved from one tower to the other. We are interested in the total number of times a block has been moved from one tower to the other tower.

This random experiment can be simulated according to the flow charts in Figure 1 and Figure 2. The first flow chart is for a programmable calculator and the second flow chart describes a computer program in BASIC.

With a program according to these flow charts successive trials can be made of the experiment with the two towers.

With such a program the following observations were obtained. Here the values m=3, n=3 were used.

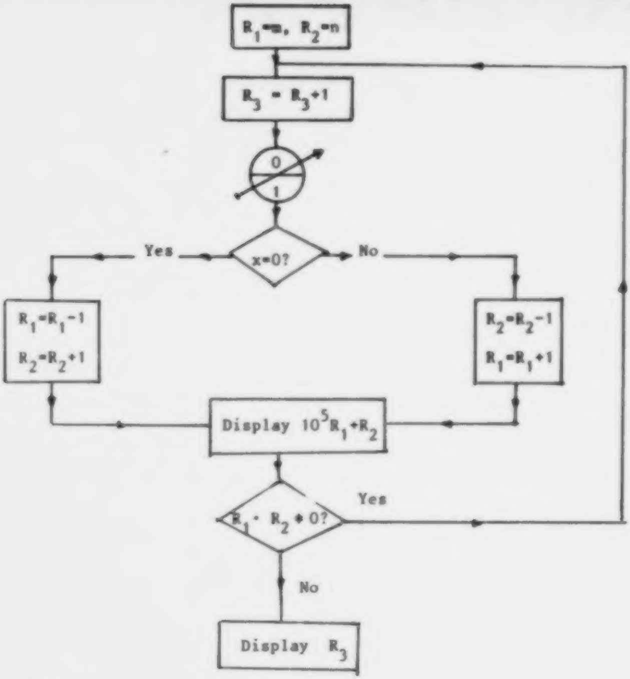


FIGURE 1

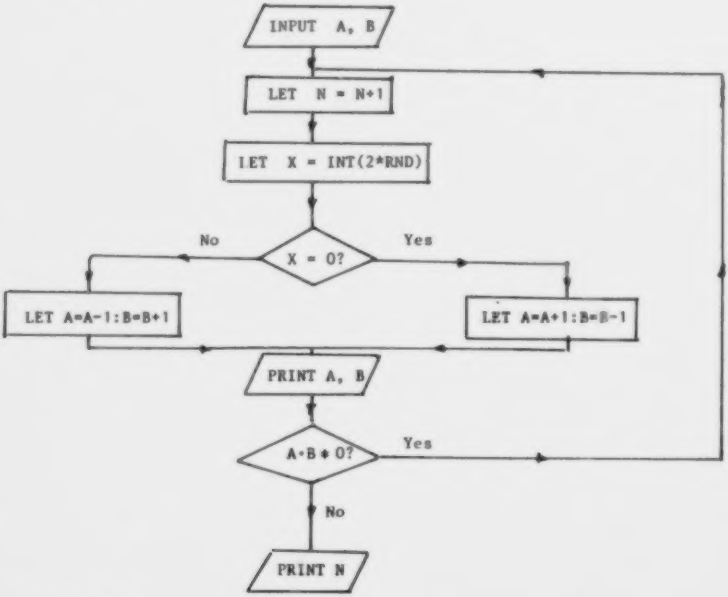


FIGURE 2

7	3	7	3	5	5	9	7	27	33
31	5	19	23	7	17	11	3	5	15
3	7	3	3	5	3	9	29	5	5
5	3	5	5	3	9	3	37	3	3
9	11	13	9	11	31	11	15	5	5
3	3	11	7	3	25	9	15	5	5
5	5	5	7	5	9	3	9	9	7
9	7	5	5	5	19	5	9	17	9
3	3	5	3	25	25	15	5	5	15
3	7	7	5	5	5	19	15	9	3

These data can be summarized by a bar graph or a box plot, see Figure 3.

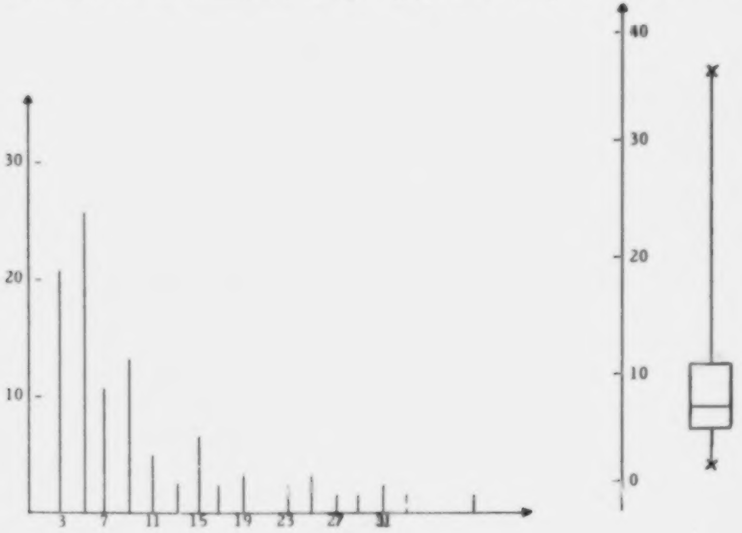


FIGURE 3

Also the mean can be calcualted. It is found to be 9.4. All this provides us with some information about the random variation of the number of moves in this case.

It is easy to modify a program according to the flow charts in Figure 1 and Figure 2 so that it will do a number of trials and calculate the average number of moves in these trials. With such a program the following averages were obtained, each time from 100 trials and with n=m=3.

10.86	8.52	9.84	9.96	9.26

Thus the mean in a large number of trials seems to be in the range 9-10. Running the program with 1000 trials gave the mean 8.838.

More insight could be obtained by doing an investigation with different values of m and n.

The following table gives the results from such an investigation where the mean in 100 trials was obtained for some different values of m and n.

n	1	2	3	4	5	6
m						
1	-	1.83	3.06	3.55	5.16	6.21
2	-	4.06	6.20	7.48	9.77	12.04
3	-	-	8.56	10.25	13.94	17.08
4	-	-	-	15.12	21.79	23.30
5	-	-	-	-	24.62	31.90
6	-	-	-	-	-	32.50

In Figure 4 these results have been graphed. Averages for $m \neq n$ have been plotted twice by interchanging m and n.

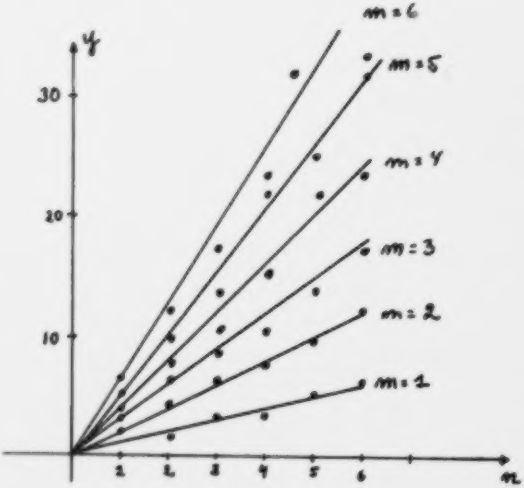


FIGURE 4

These results indicate that the mean in a large number of trials (the expectation) for given value of m is proportional to n. As a matter of fact the results indicate that the expectation is equal to the product m·n. In Figure 4 also the lines m·n have been drawn.

Is it possible to prove theoretically that the average in a large number of trials (the expectation) is m·n? Let us look on the case m=n=3. Then the random experiment can be described by the graph in Figure 5.

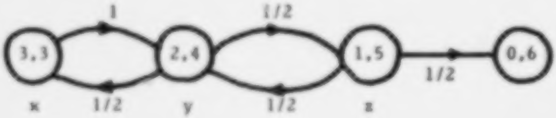


FIGURE 5

Let x, y, z be the expectations from the different states according to Figure 5. It is then easy to derive the following equations.

$$x = 1 + y$$
$$y = 1 + \tfrac{1}{2}x + \tfrac{1}{2}z$$
$$z = 1 + \tfrac{1}{2}y$$

This gives

$$x = 9 = 3.3 \qquad y = 8 = 2.4 \qquad z = 5 = 1.5$$

A similar analysis is possible in the general case.

Let us do one more thing before we leave this introductory example. It is easy to modify a program according to the flow charts in Figures 1 and 2 so that it will do successive trials and calculate the mean of the observations obtained so far after each trial. The results in Figure 6 have been obtained in this way.

The problem which has been discussed here is a classical problem in probability theory. Usually it is called the problem about "the duration of a game". In this context the problem has to do with two players, who make successive trials of a fair game. Initially they have m and n monetary units and after each trial the winner gets one such unit from the other player. After how many trials will one player be ruined?

This problem may look somewhat remote from practical interest. So let us consider a more direct practical application.

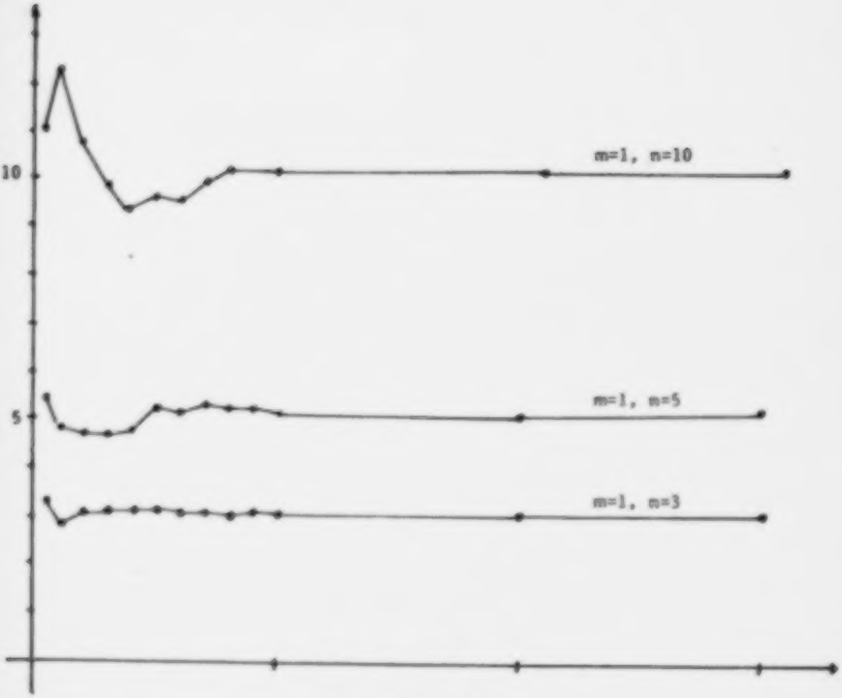


FIGURE 6

RELIABILITY AND COIN TOSSING

Consider the following random experiments dealing with the tossing of coins. First n coins are tossed simultaneously. Those which come up heads are removed and those which show tails are tossed again. The process continues until each coin has come up heads.

We are interested in the number of (group) tosses to accomplish this.

This again looks like a typical problem from a textbook in probability, charming but not dealing with a real life application. The problem, however, has an interesting application concerning the reliability of a system.

Consider a parallel system which consists of n components. The system is subject to shocks which cause the components to fail independent of each other with probability p. Each shock will remove a random number of the components like each toss in the previous problem will remove a random number of coins. We thus have the same problem if we let the coins give the outcome "head" with probability p.

This random experiment can be simulated with programs according to the flow charts in Figure 7.

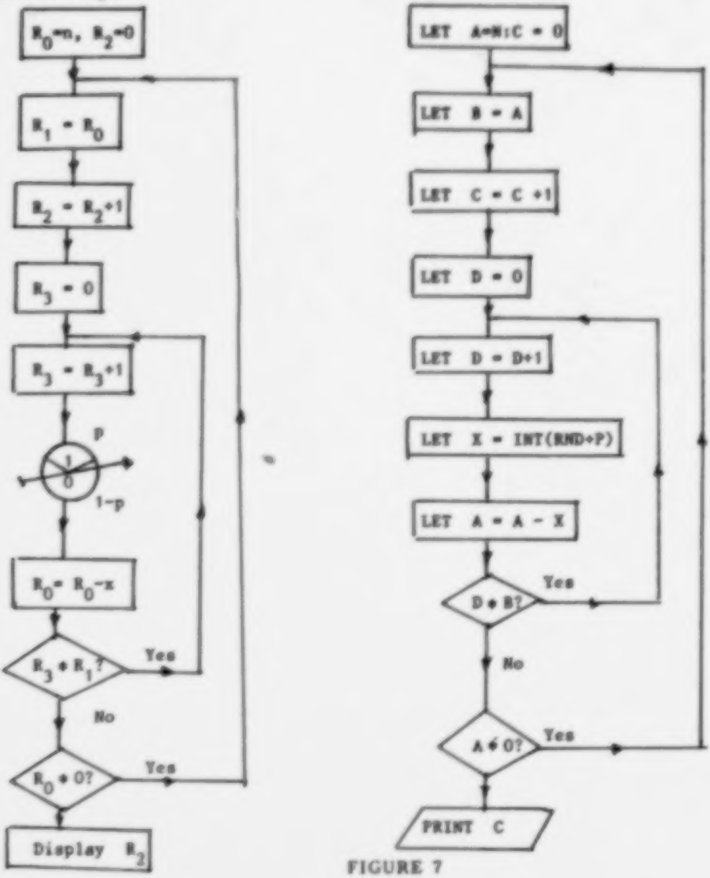


FIGURE 7

Registers and variables are used as follows in these programs.

R_0, A = number of components which have not failed.
R_1, B = number of components which are hit by shock
R_2, C = number of shocks
R_3, D = variable which counts the components in the system when hit by shock

The programs give the length of life of the system in terms of number of shocks until failure of the system.

The following lengths of life have been obtained for two specific cases with a program according to Figure 7.

n=2, p=0.3	n=5, p=0.1
4, 1, 2, 3, 5, 1, 5, 4, 1, 5	15, 8,11,26,13,17,12,21, 3,11
3,16,14, 4, 3, 1, 2, 6, 9, 2	22, 9,48,21,22,16,26,26,18,35
12, 1, 7, 3, 2, 5, 5, 1, 1, 2	21,22, 5,26,46,20,43,20,29,25
3,19, 7, 1, 2, 7, 8,10, 2, 8	16,10, 9,27,17,24,39,28,19,23
9, 3, 6, 4, 5, 4, 2, 6, 4, 8	11,17,22,14,17,12,55,16,31,19
1, 5, 1, 7, 9, 9, 3, 2, 6, 2	16,43,43,21,13,10,14,49,32,20
4, 3, 5, 4, 6, 2, 8, 6, 8, 1	16,16,17,34, 8,15,17,31,15,14
7, 5, 2, 3, 5, 1,19, 8, 1, 1	29,10, 9,40, 7,21,38,26, 5,10
4, 4, 4, 5,13, 9, 2, 3, 2, 9	11,59,28,33,15,17,28,20,16,16
6, 5, 2, 3, 6, 2, 3, 6, 2, 3	40,17,30,14,19,38,12,14,15,15

The 50 first observations in the two cases are shown in the following stem and leaf diagrams.

n=2, p=0.3			n=5, p=0.1		
0*	1231133122132112312232	(22)	0	83959	(5)
0·	455454655645464	(15)	1	5137216860791747269	(19)
0··	97778898	(8)	2	6121266126009574832	(19)
1*	20	(2)	3	591	(3)
1·	64	(2)	4	863	(3)
1··	9	(1)	5	5	(1)
		(50)			(50)

* 0,1,2,3.

· 4,5,6

·· 7,8,9

The length of life distribution is obviously very different in the two cases.

A program according to the flow charts in Figure 7 can easily be modified to give the average length of life in a number of trials. With such a program the following data have been obtained. In each case 100 trials have been made.

Average length of life in 100 trials						
	p					
n	0.05	0.10	0.15	0.20	0.25	0.30
1	21.67	10.23	6.89	5.13	4.12	3.48
2	29.24	15.20	10.27	7.13	6.39	5.10
3	41.34	16.55	11.81	8.29	6.47	5.60
4	38.75	18.64	13.77	9.58	7.64	6.13
5	44.69	22.59	14.92	10.56	8.57	6.85
6	47.71	24.59	14.83	12.06	8.83	7.39
7	53.54	26.41	16.99	12.37	9.63	7.60
8	52.90	23.96	17.86	13.17	9.66	8.13
9	55.52	25.67	18.17	12.40	10.65	8.94
10	57.94	28.39	19.45	14.10	10.96	8.85

Let us now take a closer look on the case n=2. Then the situation can be described by the diagram in Figure 8.

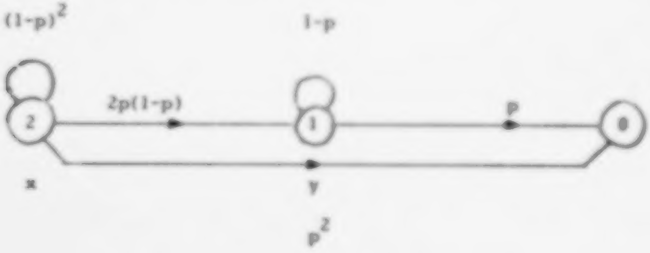


FIGURE 8

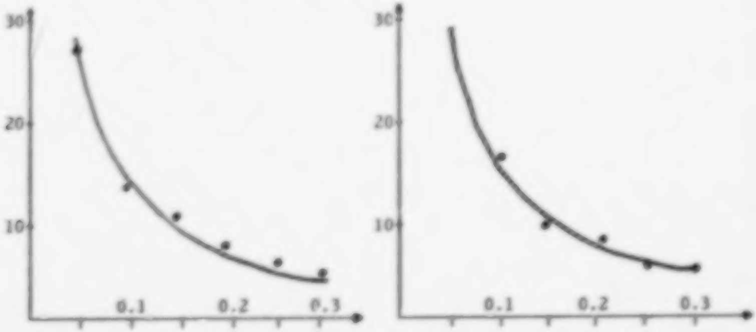
Let x and y be expectations from states according to Figure 8. Then

$$x = 1 + (1-p)^2 x + 2p(1-p)y + p^2 \cdot 0$$
$$y = 1 + (1-p)y + p \cdot 0$$

This gives $x = (3-2p)/p(2-p)$, which is the expected length of life of the system for n=2. The following table shows this expectation for p=0.05(0.05)0.30 and also means obtained in 100 and in 1000 trials.

p	Expected system life	Average in 100 trials	Average in 1000 trials
0.05	29.74	29.283	
0.10	14.73	14.07	15.158
0.15	9.73	10.88	9.483
0.20	7.22	7.68	7.231
0.25	5.72	5.81	5.624
0.30	4.71	4.73	4.728

These results are also shown in Figure 9.



Expected system life and average life in 100 trials

Expected system life and average life in 1000 trials

FIGURE 9

It is seen that already 100 trials give a very good estimate of expected system life and that 1000 trials usually (but not always) give a better estimate than 100 trials.

Let us now try to investigate the case n=5 in the same way. The situation is now illustrated by the graph in Figure 10, where, however, we have not indicated the probabilities for transitions between states.

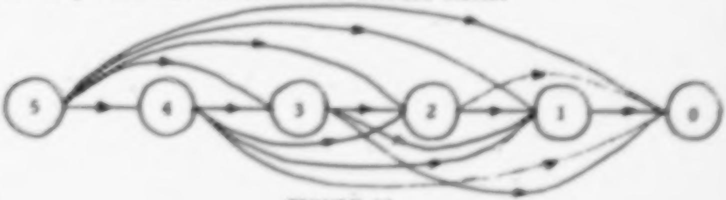
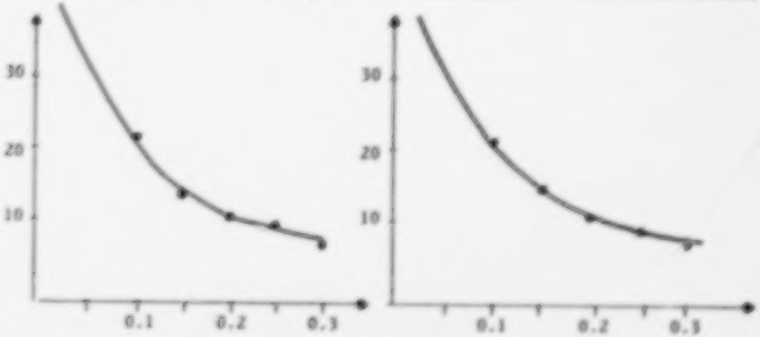


FIGURE 10

It seems to be a very laborious task to try to calculate the expectation in this case. It is, however, easy to simulate the situation with the program used in the previous case but now with n=5 instead of 2. The following data were obtained in such a way.

p	Average in 100 trials	Average in 1000 trials
0.05	42.09	45.130
0.10	22.70	22.546
0.15	14.13	14.582
0.20	10.90	10.915
0.25	8.81	8.451
0.30	6.50	6.558



Average system life in 100 trials

Average system life in 1000 trials

FIGURE 11

These averages are shown in Figure 11. We have in this case not calculated the expectation as a function of the component probability p but we can instead make an estimate by fitting a curve to the observations.

We can also combine a simulation program with a statistical data analysis program. The flow charts in Figure 12 show how programs can be written, which will calculate mean, variance and standard deviation of simulated observations of the system life. The programs can easily be modified also to give the frequency distributions.

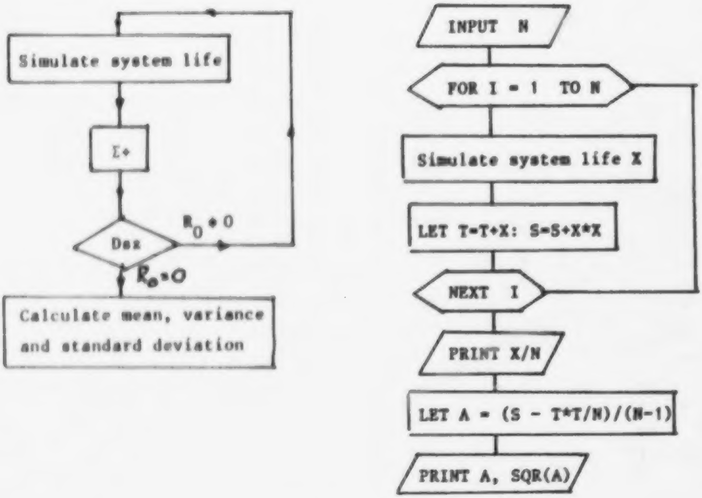


FIGURE 12

The data in the following table have been obtained with a program according to Figure 12. In all cases 100 trials have been made.

n	p	Mean	Variance	Standard deviation
2	0.05	29.7	440	21.0
2	0.10	15.4	108	10.4
2	0.15	9.41	43.6	6.61
2	0.20	7.08	26.1	5.11
2	0.25	5.66	13.9	3.73
2	0.30	4.54	7.87	2.80
5	0.05	42.9	516	22.7
5	0.10	20.8	89.6	9.46
5	0.15	15.4	78.6	8.86
5	0.20	11.0	35.0	5.91
5	0.25	8.88	19.8	4.46
5	0.30	7.13	11.6	3.41

Why is it better to estimate the expected system life with the average in 1000 trials than in 100 trials? This problem can be explored by simulating a sequence of averages and by calculating the variance for the averages. If s^2 is the variance for 100 (say) averages of n observations each, the data obtained will soon indicate that it may be suitable to calculate ns^2. The following data were obtained in this way.

Variance s^2 for 100 and 1000 means of n observations.
In all cases a system with two components and p=0.3 has been used.

100 means			1000 means		
n	s^2	ns^2	n	s^2	ns^2
1	9.39	9.39	1	9.98	9.98
2	4.08	8.16	2	5.02	10.0
3	3.50	10.5	3	3.31	9.92
4	2.67	10.7	4	2.38	9.51
5	1.88	9.40	5	2.02	10.1
6	1.55	9.30	6	1.62	9.72
7	1.42	9.94	7	1.37	9.59
8	1.47	11.81	8	1.31	10.5
9	1.14	10.3	9	1.14	10.2
10	0.981	9.81	10	1.05	10.5

The data strongly indicate that the variance of a mean of n observations is 1/n of the variance for one observation.

It should finally be remarked that it is possible to derive a closed expression for the expectation in the general case with n components or n coins; see J. Kinney [3] and L. Råde [10].

REAL DATA - SIMULATED DATA

The students will in a computer oriented course in stochastics to a large extent work with simulated data. The statistical treatment of such data is of course the same as for real data with the exception that simulated data are often generated within the computer and are not to be given from outside to the computer for statistical data analysis. But the students should also work with real data and should see some examples of how difficult it may be to distinguish between real data and simulated data. This point is discussed in Rao [4] and in Råde [6].

OTHER PROBLEMS

The following is a list of problem areas which provide lots of challenging problems which can be used in stochastics courses at the school level.

Random number generators
Test of random number generators
Simulation of probability distributions
Calculation of probability distributions
Simulation of games of chance
Stochastic geometry
Stochastic number theory
Simulation of Stochastic service systems
Simulation of storage systems
Simulation of statistical estimation procedures

CHARACTERISTICS OF A COMPUTER ORIENTED APPROACH TO STOCHASTICS

The following is a list of some advantages of a computer oriented approach to the teaching of stochastics at the school level. Hopefully these advantages have been demonstrated in the two examples above.

1) A combination of a theoretical and an experimental approach may be used.
2) Important concepts like probabilities, expectation and variance and their properties can be motivated and illustrated by appropriate experiments.
3) Work with data obtained from simulations will provide lots of opportunities for statistical data analysis.
4) Stability and variability in random experiments and statistical estimation procedures can be illustrated.

5) The students will learn about simulation, which is an important method of solving complex probability problems.

6) Simulation of random experiments can be organized as a teamwork of groups of students.

References

[1] Engel, A. (1977) Elementharmathematik von algorithmischen Standpunkt, Ernst Kleff, Stuttgart.

[2] Engel, A. (1982) Statistik auf der Schule: Ideen und Beispiele aus neuerer Zeit, Der Mathematikunterricht, 28: 1.

[3] Kinney, J. (1978) Tossing coins until all are heads, Mathematical Magazine, 51, 1.

[4] Rao, C.R. (1970) A multidisciplinary approach for teaching statistics and probability, The Teaching of Probability and Statistics: (ed. L.Råde, Almqvist & Wiksell, Uppsala.

[5] Råde, L. (1976) Probability, simulation and programmable calculators, Anwendungsorientierte Mathematik in der Sekundarstufe II (eds. (eds. W. Dörfer, R. Fischer), Verlag Johannes Heyn, Klagenfurt.

[6] Råde, L. (1981) Statistics, Simulation and Statistical Data Analysis at the School Level, Stochastik im Schulunterricht, Holder-Pichler-Tempsky, Wien.

[7] Råde, L. (1981) Random digits and the programmable calculator, Teaching Statistics and Probability, NCTM yearbook (ed. A.P.Shulte), Washington.

[8] Råde, L. (1977) Take a Chance with your calculator, Probability problems for programmable calculators, Dilithium Press.

[9] Råde, L. (1981) Algorithms for calculation and simulation of probability distributions, Department of Mathematics, Chalmers University of Technology.

[10] Råde, L. (1976) Reliability systems in random environment, Journal of Applied Probability, 13.

OPTIMUM BALANCE BETWEEN STATISTICAL THEORY AND APPLICATIONS IN TEACHING

C Radhakrishna Rao, University of Pittsburgh

There have been several round table meetings and discussion sessions on teaching of statistics organized by the International Statistical Institute (ISI) and the International Committee of Mathematical Instruction. These meetings and the deliberations of the Education Committee of the ISI at the last fifteen biennial sessions have served a useful purpose in identifying the various problems connected with the education and training of statisticians and bringing to the forefront several issues which need to be discussed in a wider forum. I am glad that the ISI has taken the responsibility in this direction to organize the first international conference on the teaching of statistics. The program committee headed by Lennart Råde deserves our congratulations for working out a comprehensive program and securing the participation of experts from all over the world. We hope that our deliberations will result in specific recommendations and, what is more important, that we will be able to set up a suitable machinery for following up on these recommendations, bringing them to the attention of educational institutions all over the world and helping in their implementation.

The program committee has placed on me a heavy responsibility in asking me to speak at one of the plenary sessions. I could not resist the temptation of accepting the invitation to participate in a conference on a subject of considerable interest to me. As Director of Research and Training at the Indian Statistical Institute for over 35 years, I had the opportunity to develop and institute a wide variety of educational and training programs in statistics at various levels:

- Pre college, undergraduate and graduate courses leading to bachelor's, master's and Ph.D. degrees.
- Applied courses for research scientists working in basic disciplines like biology, psychology, sociology and economics.
- Refresher and advanced courses for statisticians employed in government offices, businese, industry and research organizations.
- Short courses for factory workers to help in the implementation of quality control programs.
- Workshops for field workers who gather information by interviewing people or by direct observation.

I shall describe the efforts we have made in formulating these varied types of educational and training programs and in implementing them. I hope my experience will be of some use to others engaged in these endeavours.

WHITHER STATISTICS?

Much has been said about statistics as a discipline, a science and a profession. One has only to read the presidential addresses appearing in the JASA and the JRSS year after year, expressing concern about the future of statistics, lamenting the nonrecognition of statistics as a separate discipline by scientific academies, criticizing the statistical educational programs in the Universities as too theoretical, emphasizing the need to close the "vicious gap" between theoreticians and practitioners of statistics, diversification of statistics into narrow areas of specialisation, and so on. They all express the same anxiety. What is the future of statistics? How to develop statistics as a truly viable and socially meaningful science?

Whatever may be the views expressed on the status of statistics as compared to the basic disciplines, it is a fact that statistics has become an indispensable tool in all branches of human endeavour from scientific research and complex decision making to regulation of our daily lives. In his presidential address on "The Expansion of Statistics" to the Royal Statistical Society thirty years ago, R.A. Fisher (1953), described statistical science as

> ..... the peculiar aspect of human progress which gave to the twentieth century its special character; and indeed members of my present audience will know from their personal and professional experience that it is to the statistician that the present age turns for what is most essential in all its important activities. They are the "backroom boys" of every significant enterprise.

The sudden spurt in the applications of statistics which started in the second quarter of this century has led to phenomenal developments in statistical methodology such as design of experiments, sample surveys, quality control and sequential analysis. The large membership of statistical societies all over the world, and the ever increasing number of journals devoted to the theory and applications of statistics ensure us that statistical activity is a growing concern. It is growing not because statistics has reached the self generating stage like mathematics where it can create its own problems but because of its need for investigations in other areas. Unlike in other disciplines we do not study statistics to solve problems in statistics but to solve problems of other disciplines, which gives statistics its special character.

While new problems in statistics arise from its interface with other disciplines, the development of statistical methodology itself requires good knowledge of mathematics. It is interesting to note that in the table given by Box (1980) to illustrate that "interaction with reality can be very fruitful", the names of the investigators mentioned against the practical problems which led to important techniques in statistics, are those of Gauss, Bernoulli, Laplace, Wald, Fisher, etc., who are all outstanding mathematicians.

Both Fisher and Mahalanobis emphasised that progress in statistics can take place only through new applications while agreeing on the relevance and use of mathematics in developing the statistical tools. I quote from Mahalanobis (1965):

> Pure mathematics is indispensable for supplying rigorous logical foundations and for exploring the limits and refinements of statistical theory but has never given it essentially new concepts or tools. New statistical concepts and method have continually emerged only in dealing with observations in the world of reality.

The level of mathematical education necessary for a practising statistician is, of course, a different issue which has to be carefully examined.

STATISTICS AT THE HIGH SCHOOL LEVEL

Mathematics is essential and has to be taught at a young age as a method of reasoning and a discipline in rigor essential for the growing mind, and as a preparation for the study of basic sciences later in their educational career. That all sciences are in the abstract mathematics is the basic argument for teaching mathematics at an adequate level of sophistication in the high school. But study of mathematics emphasizes only one aspect of reasoning, deductive in character that of arguing from given premises or deriving the consequences from a given set of postulates. This is highly restrictive in nature and allows the students to see only one side of the coin. It is, perhaps, equally important to expose the young minds to reasoning the other way around, inductive in character of arguing from given consequences to the premises, by which new knowledge is created. While deductive reasoning was systematised through the study of mathematics more than 20 centuries ago, inductive reasoning because of its uncertain nature remained uncodified. It is only in the beginning of the present century with the emergence of statistics as a method of drawing "inference on a population from a sample" that serious attempts have been made to provide rules for inductive reasoning through appropriate statistical methodology. Thus, while mathematics provides an abstract framework for describing natural phenomena, statistics provides the concrete methodological tools in searching for a working, if not true, model for observed phenomena.

If we agree that it is important to expose the students at an early age to inductive type of reasoning as well through the study of statistics, then we have to think seriously how statistics should be introduced as a part of the high school curriculum. Should statistics be introduced as a separate subject to replace or by cutting down the syllabus of some subject that is being currently taught? Should statistics be taught to all students in the high school or only to those who wish to specialize in social and life sciences later. The two papers by Peter Holmes (1981) in the Proceedings of the Gothenburg Symposium on Teaching Statistics held in 1980 contain useful thoughts on these issues. I quote from one of the articles:

> ..... there was a growth in the numerical content of the subjects such as geography, economics and life sciences and representation was made to the examination boards (in the U.K.) that it would be much more useful for students in these courses if, in their mathematical courses they could study probability and statistics instead of theoretical mechanics. In response to this call the examination boards began to make available examinations in pure mathematics and statistics as an alternative to the more traditional pure mathematics and theoretical mechanics.

This is only to give one instance where statistics was introduced by cutting down some topics in the syllabus for mathematics, and statistics was made a compulsory subject of study only for some students. Is this the right thing to do? Some of us may not think so.

What about the contents of the course itself once it is decided to introduce statistics as a separate subject at the high school level? Naturally, one thinks of elementary probability to be taught through exercises involving permutations and combinations, probability distributions through histograms, calculation of median, mode and a few moments by using specified formulae, concepts of correlation and regression and elementary testing of hypotheses. Peter Holmes points out that over a period of years since 1959 when statistics was introduced, there was a growing dissatisfaction with these courses. This should be evident from the high school texts prepared to cover the syllabus in statistics from which the teacher is supposed to teach and from which the students are expected to memorize the formulae for answering the numerical exercises set in the examinations. The students find such a course dull and drab, and it does not serve the purpose for which a case is made for the study of statistics at an early age, viz., to inculcate the spirit of enquiry and to expose the growing minds to concepts of inductive reasoning. Peter Holmes mentions that some changes have been made in the syllabus to include project work to bring the students in contact with live data. This is desirable, although much depends on the teacher, and how the project work is conceived and the role of statistics as a method of quantitative reasoning is emphasised.

What is perhaps ideal is not to treat statistics as a separate subject but reframing the syllabus of the social and life sciences, and in natural sciences as well, injecting statistical concepts and methodology. Experiments of this nature were attempted at the Indian Statistical Institute under the guidance of J.B.S. Haldane who himself designed and taught courses in statistics at the precollege level (see Rao, 1969, 1974). This is not an easy task but worth trying.

We must remember that our aim in teaching statistics at an early age is not to impart statistical skills per se but to expose them to statistical concepts, chance mechanisms in nature and decision making under uncertainty in real situations, with examples chosen from the subjects they are studying. To arouse the interests of the students in statistics, examples given in books such as Mosteller et al (1973) and papers by Rao (1977a,b) may be discussed.

MAKING OF PROFESSIONAL STATISTICIANS

In planning for the education and training of professional statisticians we have to take into account the future demand for statisticians in different areas, the nature of problems they will be called upon to study and the skills they are expected to have. I believe the greatest demand for statisticians will be in industry and business and to some extent in the local, state and federal governments. There will be a limited demand for teachers in statistics in high schools and universities. Apart from some basic common courses in theoretical and applied statistics, different types of courses may be needed for statisticians preparing to work in different areas. In an excellent report on "Preparing statisticians for careers in industry" the ASA committee on training of statisticians in industry (with R.D. Snee as chairman) noted:

> When a statistician leaves school his or her skills and experience should include statistical knowledge, practical problem solving, consulting practice, and ability to communicate orally and in writing with non-statisticians.

This view expressed by the committee is generally applicable to the training of all types of practising statisticians. The ASA committee made a number of recommendations with special reference to the training of statisticians for industry (see ASA, 1980). It would be useful to have similar reports for training in other areas of applications of statistics.

Based on my experience as a teacher and a consultant in statistics, I suggest the following program of studies at the undergraduate and graduate levels for the training of professional statisticians. The courses are divided into three broad categories:

49

CATEGORY I Basic Theory	CATEGORY II Basic Applied	CATEGORY III Special Topics (Optional)
Probability theory	Design of experiments	Econometrics
Linear models and analysis of variance	Sample surveys	Time Series
Categorical data and chi-square tests	Computer simulation	Biometrics
Statistical inference	Cross examination of data	Technometrics
(a) hypothesis testing (b) estimation (c) decision theory		Advanced topics in (a) probability (b) inference (c) multivariate analysis

The courses in category I are meant to provide the basic knowledge in statistical concepts and theory and may be taught without relevance to particular applications. These courses should be compulsory for all statisticians. Some of the topics in these courses can be treated at different levels of mathematical sophistication depending on the mathematical equipment of the students. But knowledge of advanced calculus should be sufficient for a rigorous treatment of most of the topics.

The courses in category II will be of great value to practising statisticians in any field, although some techniques are more commonly used than the others in specific areas. The nature of instruction in these courses should be different from those in category I. In subjects like design of experiments and sample surveys, the theory can be built around practical applications. There are excellent books on these subjects which give numerous examples that may be used to provide proper motivation for developing the theory to meet different practical situations (see Des Raj (1968), Kish (1967), Murthy (1967) and Zarcovich (1966) on sampling). These subjects are ideal for introducing students to new concepts like pilot surveys, local control, cost and variance functions and uniformity trials, all of which are useful in designing optimum surveys and experiments. They also provide situations for examining problems like defining a population in terms of a frame and units for sampling, representativeness of a sample, response errors, missing values and repeated surveys, which a practising statistician in any field should be aware of.

In addition to regular instruction in theory illustrated with examples, it will be useful to assign projects to students, which involve designing of experiments and/or sample surveys, collection and analysis of data. One can think of a wide variety of projects which the students can undertake within the limited time and resources available to them while taking these courses. It will be useful if a guide book on this aspect of training can be prepared listing the titles of possible topics for project work and indicating in some detail how each project can be organised.

The possibility of arranging visits of students to experimental farms and government departments engaged in large scale surveys should be explored. Such visits would give the students an opportunity to see how the theory they are learning is actually put into practice.

In category II, I have also listed two other subjects, Computer simulation and Cross examination of data which deserve great attention. We are all familiar with computer simulation which is an invaluable aid in model building and in solving or providing an insight into problems which are mathematically intractable. Experience in this area and numerical analysis in general will be a great asset to practising statisticians. There is extensive literature on this subject, and we should try to frame a suitable syllabus for a special course on computer simulation for statisticians, giving typical examples.

Cross examination of data is a term suggested by R.A. Fisher for preliminary analysis of primary data. It literally means questioning the data to elicit relevant information before embarking on more sophisticated statistical analysis. What is an appropriate model for given data? Do the observations come from a homogeneous population? Are there outliers, recording errors, contaminating observations and other deficiencies in the data? Are the given data genuine or faked? Is there evidence that the observations are edited or censored in some way? These questions are ignored or not often raised by the users of statistical methods. Textbooks on statistics do not emphasize the importance of raising these questions and eliciting answers from data, which are useful in the choice of appropriate statistical methodology and interpretation of results. These questions are not easy to answer as there are no systematic methods developed to guide us. Cross examination of data and choice of appropriate statistical methodology have remained more in the realm of art and still depend to a large extent on the skill and experience of the statistical practitioner. See for instance the illuminating papers by Fisher (1936) and Haldane (1946) on faking of genetic data, and Mahalanobis (1933) and Majumdar and Rao (1958) on scrutiny of anthropometric data (scrutiny is an earlier term used by Mahalanobis for a preliminary look at the data). However, in recent years, attempts have been made to devise systematic methods for cross examining data, and there is enough material to cover in a separate course on the subject. I suggest the following tentative outline for the syllabus.

Syllabus for cross examination of data

- Graphical methods for display of data.
- Methods for detection of outliers and other deficiencies in data.
- Methods for examining whether data are edited, censored or faked.
- Multivariate techniques such as principal components and canonical analysis.
- Critical reviews of published papers containing statistical analysis of real data.

Books by Gnanadesikan (1977) and Tukey (1970) can provide valuable material for teaching this course. They can be supplemented by making a list of original papers such as those by Fisher, Haldane and Mahalanobis already referred to, for discussion in the class. There is no dearth of published papers containing wrong applications of statistics leading to misleading interpretations, which can be selected for critical reviews by students.

The courses under category III are optional. Those who are preparing to work in industry can take courses on special statistical methods used in industry such as quality control, acceptance sampling, experimental designs for exploring response surfaces and evolutionary operation, operations research, life testing, reliability etc., which may be broadly described as Technometrics. Those who want to work in government offices and business firms can take courses on econometric models, time series, construction of economic and social indicators, which may be broadly described as Econometrics. Those who wish to work in research wings of pharmaceutical companies, hospitals and public health departments may take courses in Biometrics which includes subjects like bio-assays, applied multivariate analysis, mathematical genetics, life tables and so on. Finally, a small number of students may wish to specialise in mathematical statistics to do research on some outstanding problems in statistical theory. They may take courses on advanced topics in probability, statistical inference and multivariate analysis.

The level of special courses in category III should be somewhat high both in its theoretical content and applied flavor. At the Indian Statistical Institute, the students taking the special courses in this category are required to spend some time as apprentices in statistical offices engaged in specified statistical activities as part of their training. Thus, those who are taking courses in technometrics are sent to industrial establishments having well developed statistical divisions. Those who are taking courses in econometrics are sent to government statistical offices, and so on.

The aim of the special courses in category III is to prepare the students to

- undertake on their own responsibility practical investigations in their special area, and
- identify new problems and do research by themselves and/or in collaboration with others for developing new statistical methodology.

THE Ph.D. PROGRAM

I have discussed the making of a professional statistician through appropriate courses of study and training at the Universities. These courses should qualify a student for a Master's degree as a mark of a high level of professional attainment. The Ph.D. program for statisticians, on the other hand, requires different considerations. The standards for such a program differ considerably from university to university within a country and between different countries. The standards for the award of the Ph.D. degree have also changed considerably over the years.

During my first visit to the United States, I asked the head of the statistics department of a well known university as to what kind of program they had in statistics. He replied with a feeling of satisfaction that it was graduate leading to a Ph.D. Then I enquired as to what they were preparing the students for. He said, for the Ph.D. degree in statistics. This was quite in contrast to my own experience while working with R.A. Fisher and at the Indian Statistical Institute. Before I went to Cambridge on an assignment to work on an anthropometric project, I had already done research work in theoretical statistics which could have earned for me a Ph.D. degree in India. However, I thought I should work with Fisher and get a Ph.D. degree from Cambridge. During my first interview with Fisher, I asked him to suggest a problem for my Ph.D. thesis. In his characteristic manner, Fisher said, "The problem must be yours and I might be able to offer you some help in solving it". I was reminded of what Poincaré said emphasising the ability to see a problem as the major step in research. "The question is what is the question". A Master's program in statistics is not worthwhile if a student who has completed the program is unable to find by himself or herself a new problem for investigation. In Cambridge, I was working on an anthropological project which involved measuring bones and stones and doing a comparative statistical analysis of observations made on different collections of skeletal material from Africa. I was aware that the statistical discriminant function for deciding between two alternatives had its origin in sexing a human mandible. Fisher was glad when I chose discrimination between several populations, which was relevant to my project, as my thesis topic. On my return to India, we set up the tradition at the Indian Statistical Institute of encouraging the Ph.D. students to choose their own problems for writing the thesis. This policy resulted in several highly original theses on topics, which were, more often, unfamiliar to the supervisors. I was surprised at what Box (1980) has mentioned:

> The student may receive verbal and nonverbal messages that tell him that the kind of Ph.D. thesis that is motivated by a scientific problem with real data is not likely to be so acceptable as a pretty but narrowly introspective mathematical exercise..... He may suspect that when his publications are assessed, those with any taint of application will be down-weighted, if not entirely discounted.

This is contrary to the very essence of research, and I hope good sense will prevail in departments where such views are held.

REFRESHER COURSES FOR PRACTITIONERS

During the last few years, the ISI has been concerned with the problems of the "growing gap between theoretical and practical statisticians" and "increasing specialisation in various subject matter areas of statistics leading to diversification of interests". I have referred to some of these problems in two papers entitled "Growing responsibilities of government statisticians" (Rao, 1975) and "Perspectives in statistics" (Rao, 1979). I believe these problems need a three pronged attack.

First, theoretical statisticians working in universities and research organisations should be encouraged to work from time to time in government statistical offices and industrial establishments. This will provide an opportunity for the academic and practising statisticians to interact and benefit from each other's knowledge and expertise. The visiting theoretical statisticians can study some of the real problems faced by the practitioners and hold seminars and workshops to discuss the possibilities of using new theoretical developments in the analysis and interpretation of data.

Second, the officers working in the government and in the industry should be given an opportunity to take refresher courses on the new developments in statistical theory. These courses should be somewhat different from the academic type given in the university departments. They should be specially designed to meet the needs of practising statisticians in different areas, with emphasis on new applications. The officers taking such courses should be asked to bring data from their offices and analyse in consultation with the teaching staff. At the Indian Statistical Institute, such special courses of three to six months duration are regularly offered to officers deputed from industry and government.

Third, there should be periodical conferences of statisticians working in specific areas to discuss current developments. It will serve a useful purpose if such conferences are centered round a few problems on which key note papers by both theoreticians and practitioners are invited and discussed.

COURSES FOR RESEARCH SCIENTISTS

At present, the method of providing instruction in statistics to students specializing in other disciplines at the college level is far from satisfactory. In some universities all the courses in statistics are centralized and are offered by the department of statistics, which announces a list of courses ranging from elementary statistics to highly specialized topics. Students from other disciplines can take one or more of these courses given by the faculty of the statistics department. In other universities, each department for a particular discipline formulates its own courses in statistics and has a faculty member to give these courses to their students. Both practices have some advantages and disadvantages.

It must be noted that the courses in statistics for students in other disciplines can have only a limited objective:

- Make them aware of the usefulness of the statistical approach to problem solving.
- Demonstrate errors in measurements, differences between observers and instruments, and scrutiny of data.
- Graphical display of data and summarization of data in terms of quantiles, mean, mode and lower order moments.
- Principles of design of experiments, difficulties in interpretation of historical data.
- Preparation of schedules for recording observations.
- Tests based on mean values and proportions. Concept of the standard error.
- Use of computers in data analysis.

The courses should have a strong applied bias with examples chosen from the fields in which the students are specializing and emphasis on the need for properly defining the measurements, making observations and ensuring comparability of results. The students should be made aware of the literature in their disciplines on investigations which use more advanced statistical methodology for interpretation of results. The need to consult a competent statistician rather than blindly using computer programs in such cases may be emphasized.

The elementary courses offered by the statistics departments do not generally meet the objectives listed above, and further such courses are assigned to inexperienced staff members or teaching assistants who can only teach from cookbooks. There may be a case for specialists in each individual discipline with a good knowledge of statistics to give such courses to students specializing in that particular discipline.

THE TRAINING OF TEACHERS

The teacher is an important component in all educational and training programs. It is not enough to have an impressive sounding syllabus. More important is the availability of good teachers who have the necessary knowledge and experience to present the different topics of the syllabus in an interesting way and arouse the enthusiasm of the students, especially at the high school level. Along with our deliberations on the objectives and contents of the courses at different levels we should also consider what steps can be taken to ensure a high quality of teaching. I suggest the following for further discussion:

- Preparation of teaching manuals and audio-visual aides along with the syllabus.
- Workshops, short courses and discussion meetings for teachers.
- Encouraging teachers to take up consultation and acquire experience in solving practical problems.

It will be generally useful for teachers and research workers in statistics to get acquainted with the subject matter of a basic discipline in some depth to be able to work by themselves and/or effectively collaborate with workers in the discipline on research projects involving a quantitative approach.

IN CONCLUSION

The object of introducing statistics in the high school curriculum is not to provide them with a kit of statistical tools and a standard set of prescriptions for use in practice. It is not for producing statisticians at lower levels, if at all such a category has a meaning. It is for inculcating a spirit of enquiry and exposing the young and growing minds to inductive reasoning as complementary to deductive reasoning for advancement of knowledge. This is best achieved by not teaching statistics as a separate subject in the high school but by injecting statistical ideas and quantitative approach in the syllabus for social and life sciences.

The curriculum at the undergraduate and graduate levels should be designed to meet the future demand for statisticians at the level of practitioners in special areas of application (such as in government, industry, business, etc.), general consultants, research workers in statistical theory and competent teachers. Suitable syllabi for these purposes are outlined, providing a proper balance between theory and applications.

Refresher courses for practitioners of statistics is important in order to keep them abreast with latest developments in statistical theory and refinements in statistical methodology. This will solve, to some extent, the problem of "transfer of methodology between academic and practising statisticians".

Courses in statistics for research workers in other disciplines should emphasize the quantitative approach in problem solving, and cover the use of standard statistical methodology.

Finally, there is a need to produce teachers of statistics with a strong theoretical background and experience in the application of statistics in at least one basic discipline.

References

[1] American Statistical Association (1980) Preparing statisticians for careers in industry: Report of the ASA committee on training of statisticians for industry. The American Statistician, 34, 65-75.

[2] Box, G.E.P. (1980) The American Statistician, 34, 76-77.

[3] Des Raj (1968) Sampling Theory, Macgraw Hill.

[4] Fisher, R.A. (1936) Has Mendel's work been rediscovered? Annals of Science, 1, 115-137.

[5] Fisher, R.A. (1953) The expansion of statistics. J.Roy.Statist.Soc., A, 116, 1-6.

[6] Gnanadesikan, R. (1977) Methods for Data Analysis of Multivariate Observations, John Wiley.

[7] Haldane, J.B.S. (1946) The faking of genetic results, Eureka, 6, 21-24.

[8] Holmes, Peter (1981) Why teach statistics and what statistics should be taught to pupils aged 11-16; Using project work in teaching statistics to students aged 16-18. Proc.Gothenburg Symposium on Teaching Statistics (Editor, L. Råde).

[9] Kish, L. (1967) Survey Sampling, John Wiley.

[10] Mahalanobis, P.C. (1933) A revision of Risley's anthropometric data relating to the tribes and castes of Bengal. Sankhyā, 1, 76-105.

[11] Mahalanobis, P.C. (1965) Statistics as a key technology. The American Statistician, 19, 43-46.

[12] Majumdar, D.N. and Rao, C. Radhakrishna (1958) Bengal anthropometric survey, 1945 - a statistical study. Sankhyā, 19, 201-408.

[13] Mosteller, F., Kruskal, W.H., Link, R.F., Pieters, R.S. and Rising, G.R. (eds.) (1973) Statistics by Examples, Addison-Wiley.

[14] Murthy, M.N. (1967) Sampling Theory and Methods, Stat. Publishing Soc., Calcutta.

[15] Rao, C. Radhakrishna (1969) A multidisciplinary approach for teaching statistics and probability. Sankhyā, B, 31, 321-340.

[16] Rao, C. Radhakrishna (1971) Data, analysis and statistical thinking. Chapter 17 (pp.383-392) in Econonic and Social Development, Vora and Co., India.

[17] Rao, C. Radhakrishna (1974) Teaching of statistics at the secondary level - an interdisciplinary approach. Statistics at the School Level, 121-140, Almqvist and Wiksell.

[18] Rao, C. Radhakrishna (1975) Growing responsibilities of government statisticians. Occasional Papers, No.4, Asian Institute of Statistics, Tokyo.

[19] Rao, C. Radhakrishna (1977a) A natural example of weighted distributions - a classroom exercise. The American Statistician, 31, 24-26.

[20] Rao, C. Radhakrishna (1977b) Statistics in search of truth. Science Today, 27-36.

[21] Rao, C. Radhakrishna (1979) Perspectives in Statistics, Sankhyā B, 41, 129-137.

[22] Tukey, J. (1970) Exploratory Data Analysis, Addison Wesley.

[23] Zarkovich, S.S. (1966) Quality of Statistical Data, FAO, Rome.

APPENDIX

During the last few days, I had the opportunity to listen to various speakers at the conference on topics such as: Why teach statistics in high school? What topics in statistics should be taught in high schools and in colleges? How should statistics be taught at different levels? In this appendix, I shall comment on some of the views expressed by the previous speakers, and explain the basis of the suggestions made in section 3 of my paper on "The Making of Professional Statisticians".

At the high school level, the question of how to teach statistics is extremely important and has been the subject of considerable discussion during the last few years. But at the graduate and postgraduate levels, the question of what to teach is far more important than how to teach. In order to decide on a suitable course of study for qualifying as a professional statistician, we should recognize three main aspects of statistics as a tool in applied research, Data, Methodology and Quantitative thinking, as detailed in the following chart.

Traditionally, the term statistics is used to denote numbers, each of which is a well defined measurement made on an individual or more generally, on unit of a population (real or hypothetical). Each number has a meaning in the sense that it conveys information on the particular unit which provided the measurement. Often, we are not interested in individual units, but would

THREE ASPECTS OF STATISTICS

DATA	METHODOLOGY	QUANTITATIVE THINKING
(1) Observations (2) Mechanism Generating Observations (3) Prior Information about the Population	A. COLLECTION OF DATA (1) Sample Surveys (2) Design of Experiments B. THEORIES OF INFERENCE (1) Estimation (2) Hypothesis Testing (3) Decision Making	(1) Trouble Shooting (2) Formulation of New Problems

like to have some information about the population as a whole, comprising the different units. If all the units could be measured, then any information about the population would be known. If only some units are measured, then the inference on the population would be uncertain. The term statistics, as it is used now, stands for the whole process of making inference under uncertainty, which consists of collecting observations, analysing them, (taking into account the mechanism of generating observations and any prior information available about the population), and interpreting the results. The observations, the mechanism generating them and the prior information constitute the data. The techniques of acquiring data in an efficient way (i.e., to provide maximum information at given cost through well designed experiments and sample surveys) and answering specific questions (like testing given hypotheses and estimating specified parameters) constitute the statistical methodology. Usually, data contain more information than what is required for answering specific questions. The process of extracting such information as a guide to future investigations may be called quantitative thinking. Thus, data analysis has a dual purpose:

DATA ANALYSIS = ANSWERING SPECIFIC QUESTIONS + RAISING QUESTIONS FOR FURTHER ENQUIRY

Section 3 of the paper "The Making of Professional Statisticians" contains details of compulsory and optional subjects to be taught at graduate and postgraduate levels. Some additional remarks on teaching are as follows:

- There are considerable controversies on theories of inference and methods of data analysis, which should be brought to the attention of the students to make them aware of the limitations of statistical theory, as developed up till now.
- The importance of well documented and accurate data as prerequisite for correct inference should be emphasized.

- There is often more information in data than what can be extracted by methodology designed to answer specific questions. The method of extracting such information has no specific rules and remains as an art depending on the skill and experience of the statistician. Discussion of numerous live examples from applied papers emphasizing the three aspects of statistics in research work would be extremely useful.

This work is sponsored by the Air Force office of Scientific Research under Contract F49629-82-K-001. Reproduction in whole or in part is permitted for any purpose of the United States Government.

THE RELEVANCE OF STATISTICAL TRAINING

J Gani, University of Kentucky

INTRODUCTION

It is in the nature of every professional to regard his calling as indispensable to the common good; practising statisticians and teachers of statistics in schools and universities are no exception to this rule. Yet, in the face of competing claims for the importance of their areas from professionals in agriculture, the arts and humanities, business and administration, the law, the biological and medical sciences, the physical sciences, the social sciences and technology, statisticians may experience an occasional twinge of doubt as to the relevance and standing of their subject.

It is my intention to demonstrate in a very pragmatic way that statistics is a subject of vital importance; it enters substantially into the quantitative content of many (if not all) other professional studies. Statistical information is pervasive: a detailed analysis of The New York Times newspaper for Saturday 22 May 1982 will leave us in little doubt of this fact. The intelligent reading of such a newspaper, or of any other current information material, therefore makes a basic grounding in statistics essential for all citizens. This premise leads very naturally to a discussion of what a common core of statistical training should consist of, and how it might best be imparted to students in schools and colleges throughout the world.

At this stage we are faced with the question of whether basic statistical training can be culture-invariant. Do the citizens of a developed country require the same statistical knowledge as those of a developing nation? Is a grasp of demographic or agricultural statistics, which may be essential to a Nigerian, radically different in kind from an understanding of economic or industrial statistics, which may be more relevant to a Japanese? It may be desirable in theory for all of us to know the same elements of statistics, but in practice our training will depend on the educationl resources available in our different countries. Even if a unviersal course in basic statistics were agreed upon, the different circumstances of each nation would dictate the use of different local examples and illustrations. There would clearly be little point in using stock market statistics as an example of a time series trend in the USSR or illustrating the concept of variability with industrial quality control data in Upper Volta.

The picture that emerges from our study of The New York Times is one of considerable reliance on statistics, in the sense of numerical data, its graphical description, and its comparative evaluation. The estimative and inferential roles of statistics still remain fairly rudimentary at the layman's level of need. Yet the world-wide spread of scientific education and the growing use of home mini-computers in the developed countries make it increasingly important that a modicum of statistics be taught to every student in our schools and colleges. It has been one of the functions of this International Conference on the Teaching of Statistics to consider how this can best be done.

THE NEW YORK TIMES OF 22 MAY 1982

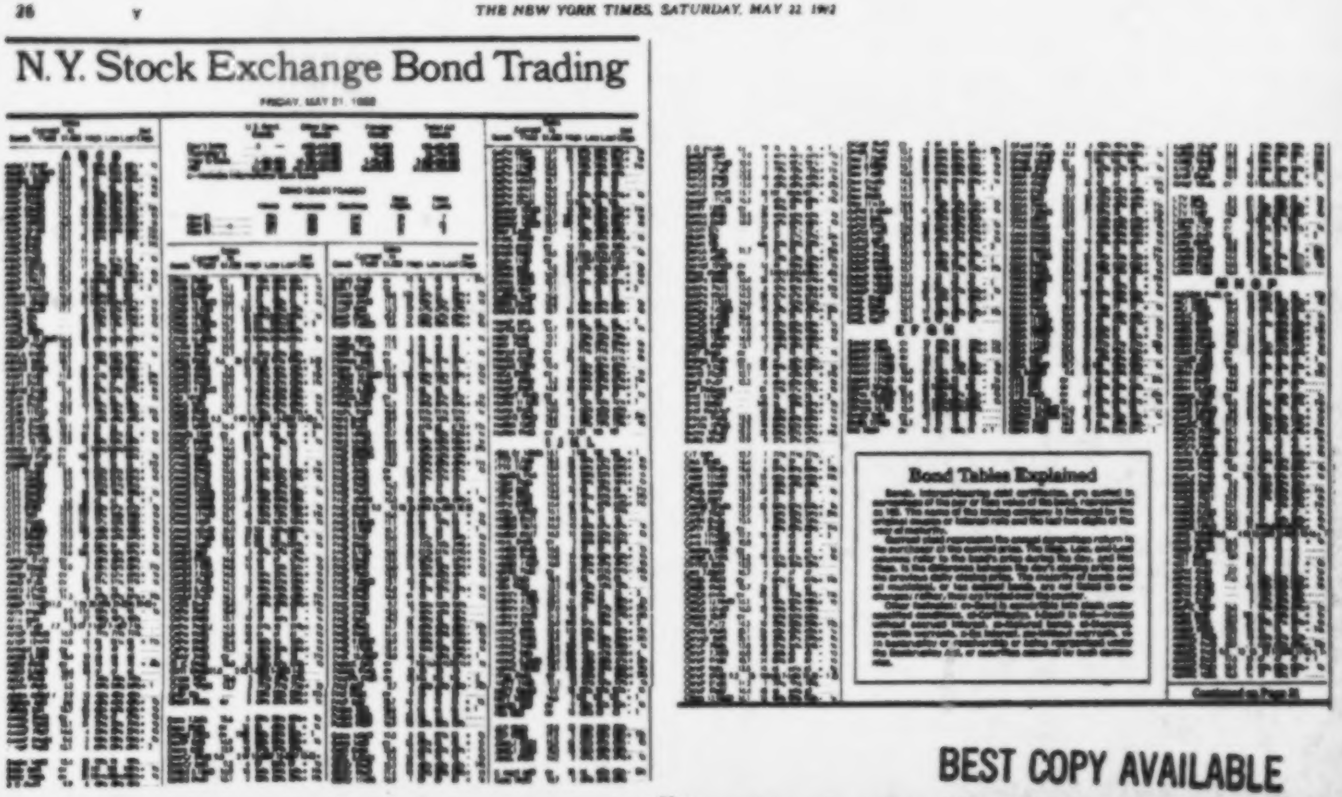
To establish the importance of statistics in our daily life, I thought it useful to analyze the contents of a newspaper of international standing, with a wide popular circulation. The New York Times, a national US daily, which I purchased in Minneapolis, seemed ideally suited for the purpose. On 22 May 1982 it consisted of 34 pages, 3 of which were advertisements, while 31 contained articles and photographs of the news of the day, together with assorted items of information such as the composition of a US Senate roll-call, tables of financial data, sports results, death notices and weather records. I estimated that the proportion of this information which contained statistical material of some kind was 32.6% of the total available space of 31 pages, or just under one third of the paper; one would have difficulty in denying the relevance of the subject in the face of such evidence. Let us examine some of the material in greater detail.

Financial Reports

The biggest single part of the newspaper relying heavily on statistics was the financial section in the 12 page "Business Day". Statistical material in this totalled 6.25 pages, or 20.2% of the newspaper's information content: it covered:

1. N.Y. Stock Exchange Bond Trading
2. Treasury Bills, Bonds and Notes
3. Money Market Mutual funds
4. Federal Reserve
5. New York Stock Exchange Issues
6. New York Stock Exchange
7. Futures Prices
8. Cash Prices (Foods, Grain and Feeds, Fats and Oils, Metals, Textiles and Fibers, Miscellaneous and Petroleum)
9. Foreign Exchange (Lstg 1 = US \$1.80, A\$ 1 = US \$1.06, Yen 1 = US \$0.004213, Sw Fr 1 = US \$0.5102)
10. World Gold (US \$335.75 per troy oz)
11. Over-the-counter Quotations
12. Trading in Stock Options
13. American Stock Exchange
14. Other US and Foreign Stock Exchanges
15. Corporate Earnings
16. Foreign Stock Index

FIGURE 1



26 V THE NEW YORK TIMES, SATURDAY, MAY 22, 1982

N.Y. Stock Exchange Bond Trading

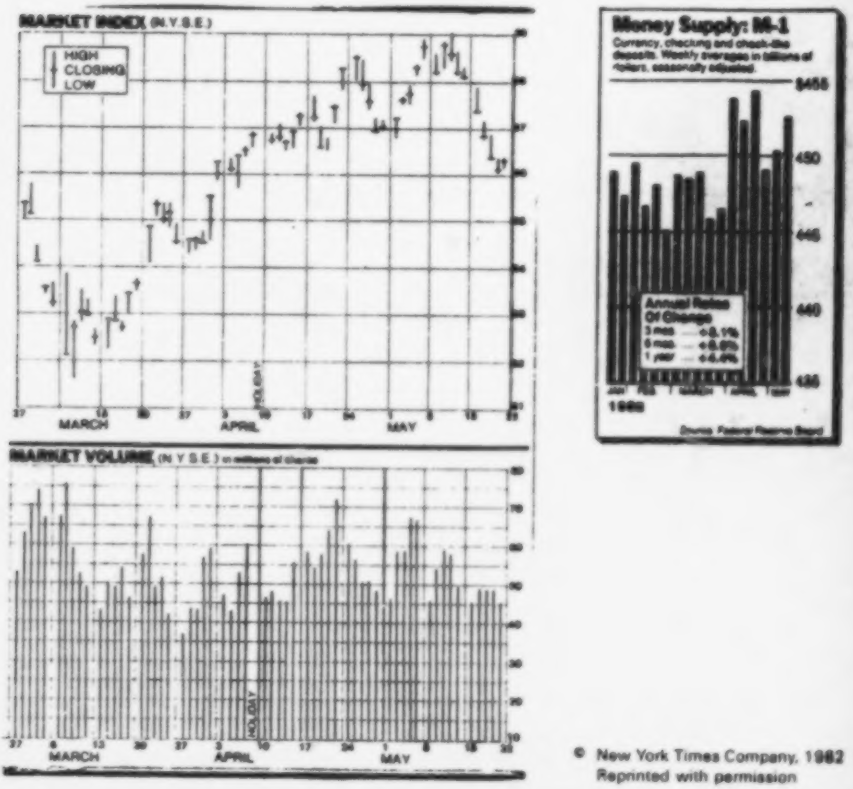
FRIDAY, MAY 21, 1982

Bond Tables Explained

BEST COPY AVAILABLE

Typical information in this section consisted of numerical data, together with its range of variation; for example in item No.3, assets in the Fidelity Tax Exempt Money market Mutual Fund maturing in 62 days, were yielding 7.47%, down 0.02% from the previous week. Or in item No.1 (see Figure 1) $10,000 worth of Aetna Life Bonds maturing in 2007 and currently yielding 13.5%, were sold at 60.25 percent, 1/2 percent above the previous day's closing price. There were also time series graphs for the market index and market volume of the New York Exchange and for the size of the money supply in the US since January 1982 (Figure 2). While none of this financial statistics is particularly complicated, it nevertheless requires both training and familiarity for its understanding.

FIGURE 2



© New York Times Company, 1982
Reprinted with permission

BEST COPY AVAILABLE

Economic Problems

The second largest section in the newspaper involving statistical information could be classified under the heading of economic problems, both US and international. This took up 4.4% of the paper's information content: the 12 articles in it covered such diverse topics as

17. Senators Prepare to Vote on Budget with $784 Billion
18. Consumer Prices Up Only 0.2%; Energy Costs Off Sharply in April
19. 252 Pieces of Federal Land Picked for Possible Sale, Watt Says
20. Living Costs Up Slightly in the New York Region
21. Furniture Freeze Violator
22. Money Supply Up $2.3 Billion
23. Orders for Durables Dropped 3.9% in April
24. Continental to Buy STSC
25. Harvester Statement
26. Stocks Halt Slide With Dow Up 3.2
27. Dominicans Fear US Sugar Quotas
28. OPEC Rein on Output Extended: Group to keep Oil Price at $34 a Barrel

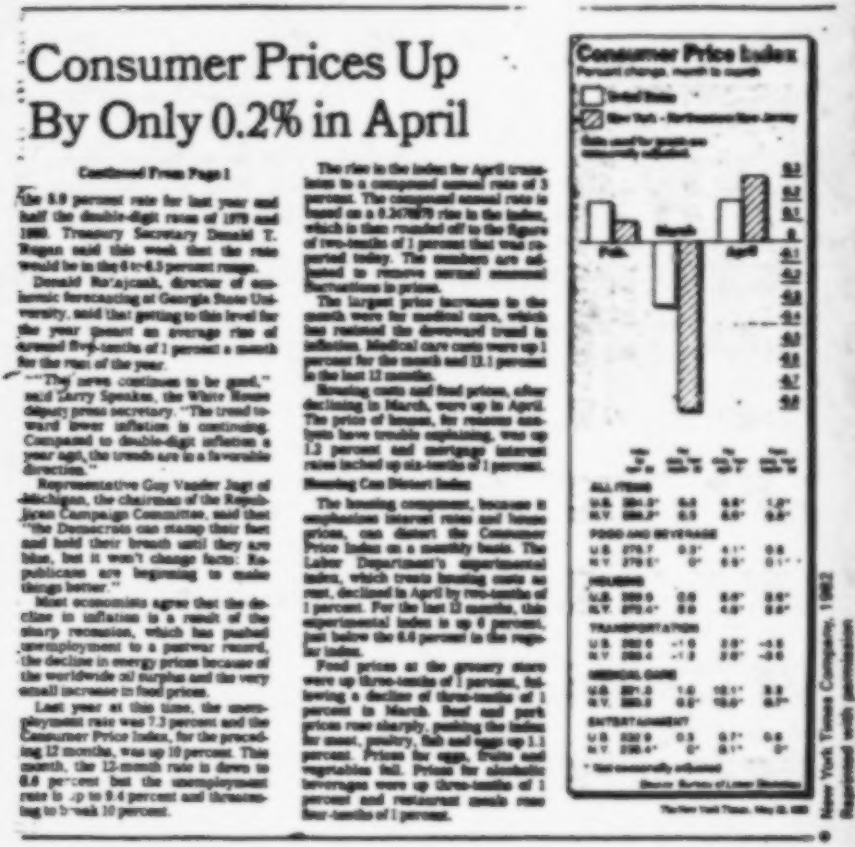
To give the flavour of these articles, and of their statistical contents let me select two among them for analysis: one of them relates to the US economy (item No.18) and the other to its effect on the Dominican Republic (item No.27).

In article No.18, it was reported that the Consumer Price Index in the US had declined by 0.3% in March 1982 (see Figure 3) for the first time in almost 17 years. Although there had been an increase of 0.2% in the Index in April this continued to reflect the slowing of inflation in the US over the past 7 months. The total increase in the Index over the last 12 months was 6.6%, the lowest since 1978; the expected rate of increase in 1982 was 6-6.5% as against 8.9% in 1981. However, at the same time in 1981, when the Index had increased by 10% over the preceding 12 months, the unemployment rate was 7.3%, whereas the current Index increase of 6.6% was associated with an unemployment rate of 9.4%, threatening to break 10%. A detailed breakdown of costs for medical expenses, housing, food, alcohol, and restaurant meals was given, together with a graph of the Index for February, March and April 1982, comparing its value for the whole of the US with that for New York and Northeastern New Jersey.

The statistical arguments here are somewhat more complex: we are dealing with rates of change in the consumer Price Index, and related changes in unemployment. Readers are now expected to understand the economic terminology, the basic statistical data, and the concept of rate of change which leads to the conclusion that the rate of inflation is gradually slowing down. In addition, they must be aware of the possible random variation of the percentage changes

BEST COPY AVAILABLE

FIGURE 3


Consumer Price Index
United States
New York - Northeastern New Jersey
Feb.
March
April

Consumer Prices Up By Only 0.2% in April

Continued From Page 1

the 8.9 percent rate for last year and half the double-digit rates of 1979 and 1980. Treasury Secretary Donald T. Regan said this week that the rate would be in the 6 or 6.5 percent range.

Donald Ratajczak, director of economic forecasting at Georgia State University, said that getting to this level for the year meant an average rise of around five-tenths of 1 percent a month for the rest of the year.

"The news continues to be good," said Larry Speakes, the White House deputy press secretary. "The trend toward lower inflation is continuing. Compared to double-digit inflation a year ago, the trends are in a favorable direction."

Representative Guy Vander Jagt of Michigan, the chairman of the Republican Campaign Committee, said that "the Democrats can stamp their feet and hold their breath until they are blue, but it won't change facts: Republicans are beginning to make things better."

Most economists agree that the decline in inflation is a result of the sharp recession, which has pushed unemployment to a postwar record, the decline in energy prices because of the worldwide oil surplus and the very small increase in food prices.

Last year at this time, the unemployment rate was 7.3 percent and the Consumer Price Index, for the preceding 12 months, was up 10 percent. This month, the 12-month rate is down to 6.6 percent but the unemployment rate is up to 9.4 percent and threatening to break 10 percent.

The rise in the index for April translates to a compound annual rate of 3 percent. The compound annual rate is based on a 0.247879 rise in the index, which is then rounded off to the figure of two-tenths of 1 percent that was reported today. The numbers are adjusted to remove normal seasonal fluctuations in prices.

The largest price increases in the month were for medical care, which has resisted the downward trend in inflation. Medical care costs were up 1 percent for the month and 12.1 percent in the last 12 months.

Housing costs and food prices, after declining in March, were up in April. The price of houses, for reasons analysts have trouble explaining, was up 1.2 percent and mortgage interest rates inched up six-tenths of 1 percent.

Housing Can Distort Index

The housing component, because it emphasizes interest rates and house prices, can distort the Consumer Price Index on a monthly basis. The Labor Department's experimental index, which treats housing costs as rent, declined in April by two-tenths of 1 percent. For the last 12 months, this experimental index is up 6 percent, just below the 6.6 percent in the regular index.

Food prices at the grocery store were up three-tenths of 1 percent, following a decline of three-tenths of 1 percent in March. Beef and pork prices rose sharply, pushing the index for meat, poultry, fish and eggs up 1.1 percent. Prices for eggs, fruits and vegetables fell. Prices for alcoholic beverages were up three-tenths of 1 percent and restaurant meals rose four-tenths of 1 percent.

New York Times Company, 1982
Reprinted with permission

in the CPI, which may mean that there is no significant difference in the February/March/April figures. The article also raises the possibility of statistical correlation between unemployment and inflation.

In article No.27 we are launched into the international arena . To help domestic sugar producers the USA is proposing a quota on sugar imports into the country, and the Dominican Republic is to get nearly 18% of this quota. However, this could still cut its present exports of sugar to the USA by one half, causing a decline of $209 million in revenue. Since 100,000 people are employed in the Dominican sugar industry, this could cause further unemployment in a country with a current unemployment rate of 30%. Because more than 60% of the sugar is produced and exported directly by the government, each pound shipped at a loss creates a larger deficit and hence higher inflation in the country. It costs the Dominican government 17 cents to produce one pound for export, while the private company Gulf and Western Industries can do it for 12 cents. The government, unlike the company, has not been able to modernise its equipment. The present world market price is just over 8 cents a pound; at current market rates, the Dominican Republic would therefore lose between 6-9 cents per pound. The Central Bank estimates that 1.13 million metric tons of sugar will be exported this year. The Dominican government has been hard pressed financially, and has not been able to find the capital to invest in a change of its land usage from sugar to alternative cash crops.

Once again readers need to understand some economic concepts: quotas, government subsidies, inflation, unemployment. They must also appreciate the comparative statistics of state and private production, and the rationale of investment in agricultural and technical improvements. There is, in addition, the possibility of a misuse of statistics; of the 100,000 employed in the sugar industry, it is not made clear how many are employed in government, and how many in private industry.

Social Issues

The third largest section containing statistical information was concerned with social isues, exclusively in the USA. This took up 3.5% of the paper's information content: the 7 articles in it were

29. Disability Cutoffs Assailed by Panel
30. 107 Seized in "Smoke Shop" Raids Arraigned on Charges in New York
31. Drops in Diseases Traced to Policy
32. Is Court the Right Place for Wayward Youths?
33. Lower Rises Seen for Some Rentals
34. For the Devotees of Liverwurst, More Good News
35. Caution on College-Cost Advisory Services

These include discussions on social security, criminological, epidemiological, housing, education and (amusingly enough) gastronomic problems. We shall again select two of these for analysis.

In article No.29, Members of Congress criticised the Social Security Administration for the way in which it was removing people "with callous indifference" from the disability insurance program. The US government pays more than $17 billion a year in disability benefits to 4.5 million disabled workers and their families, the average benefit being $419 a month. In the last 15 months there has been evidence of a rather hard drive by the federal government to remove many people entitled to benefits from the disability rolls. In the first 13 months of accelerated reviews, 436,000 cases were examined, and 174,000 or 40% of them were ruled ineligible. The article requires an understanding of social science terminology, a grasp of statistical averages, and an appreciation of whether ineligibility is a likely event in such a review, or whether it is significantly different from the previous rejection rate.

Article No.31 was concerned with the New York City policy of allowing only immunized students to enter the public school system; it was claimed that this policy had rid the city of measles, and reduced the incidence of rubella and mumps. In September, when the policy was put into effect 16% of the city's 907,000 public school students had not been immunized; 67,000 had been initially excluded at the end of September, but 47,000 of these were permitted to return in the following days if they produced the required parental consent for immunization. Only 3,163 were still being excluded because they lacked these, or around 0.35% of the total number.

The policy appeared to have been successful. Only 11 cases of measles had been reported in the first five months of 1982, as against 859 in 1979, 1,210 in 1980, and 106 in 1981. The Deputy Mayor had stated that no cases of rubella and only 30 cases of mumps had been reported. But the Director of the Health Department's immunization program reported that there had been 21 cases of rubella in the first 5 months, compared with 196 in the same period in 1979, and 30 of mumps as compared with 67 in the same period in 1975. It was clear from the article that there was a discrepancy in the Deputy Mayor's and the Health Department's reports; notwithstanding these, the program does appear to have had a positive effect.

Here, the reader needs to understand basic ideas of significance; are the results quoted truly indicative that the program has been successful? Or could the results be random? In this case the figures are fairly clear: immunization has indeed proved to be successful.

Political Arithmetic

Political issues involving statistical data, all of them international, occupied 1.7% of the information content of the paper. The 3 articles in this section were:

36. Britain Seizes a Beach Head in Falklands; Reports Downing 17 Planes, Had 5 Ships Hit.

37. State Votes Hurt Mrs. Gandhi

38. In the Namibia Talks, a New Approach.

Article 36 was full of statistics of planes lost, ships sunk, and numbers killed and wounded by the combatants. The statistical data were essentially descriptive, but there were serious discrepancies between the Argentine and the British figures. No advanced statistical concepts were necessary to understand the article.

Perhaps the most interesting of the three articles was item No.38 about the five Western Nations trying to mediate between South Africa and the South West Africa People's Organisation (SWAPO) to negotiate an independence plan for Namibia. The article in fact contains no statistical data at all, but raises very important statistical questions. It deals with a most difficult problem of voting formulae, and how these reflect the composition of the electorate. A complex voting formula has been proposed, mixing a proportional representation system with direct election on the basis of defined constituencies. South Africa was insisting on the mixed voting system, in which each ballot would, in effect, be counted twice. This would ensure that right-wing whites supporting the National Party would be represented on the Assembly drafting the constitution of an independent Namibia. Since whites are almost everywhere outnumbered by blacks, only the proportional representation system would achieve this. But this would equally favour SWAPO, whose support is heavily concentrated in the North of the country. Since South Africa's other aim is to see that insurgents fall short of the two thirds vote necessary to control the Assembly, it also favours direct voting to dilute SWAPO's support.

This article brings up some very complex statistical issues, though we can understand their basic principles well enough. We all know that the British system of voluntary voting, with the election going to the candidate with the majority vote, is different from the Australian one, with compulsory voting and the preferential transfer of votes from a losing candidate to one still in the race. The article describes the manipulations leading to the political advantage of one or other party. Of such political arithmetic is the government of countries made up, and it requires quite a sophisticated statistical understanding to grasp its finer points.

Sporting Records

These records occupied 1.7% of the paper and listed the detailed statistics of baseball games played throughout the USA, the tennis scores in the Italian Open at Rome and the German Open at West Berlin, and the seeding of both men (1. Jimmy Connors, USA) and women players (1. Chris Evert-Lloyd, USA) at the French Open at Paris due to be played on Monday 24 May (see Figure 4).

Golf scores for the Atlanta Classic played on Thursday 20 May were also listed as were the soccer results for various US soccer leagues. Such statistics are essentially descriptive, but could certainly serve as predictive data for the betting man.

Weather Information

Lastly, the weather took up 1.1% of the paper, a full 1.8 columns of the back page, over one third of it statistical in type (see Figure 5). On that day the forecast for Washington was "rain likely", temperature 73-61°F and for Minneapolis "mostly cloudy", 66-48°F. The temperatures and weather conditions on the previous day in London at noon were "cloudy", 64°F and in Buenos Aires at 9.00 a.m. "cloudy", 41°F. Once again these statistics were essentially descriptive; the forecasts and synoptic chart, while relying on complex numerical and statistical methods, were not accompanied by technical details.

Let us now summarize our survey. The contents of the New York Times of 22 May 1982, which involved at least some statistical material, can now be broken up into the following categories:

Category	Percentage space of all paper
Financial Reports	20.2%
Economic Problems	4.4%
Social Issues	3.5%
Political Arithmetic	1.7%
Sporting Records	1.7%
Weather Information	1.1%
TOTAL	32.6%

None of these categories involved difficult statistics, though a deeper scrutiny of some of the articles would inevitably lead to statistical complexities. Statistical concepts arose mostly in the financial, economic and social fields, with the recreational (sporting records) and the scientific fields (weather information) forming only a small proportion of the total. One may conclude from this that since all men are social creatures, it is most likely that they will encounter statistics in the economic and social contexts. It is clear that with nearly one third of the newspaper involving statistical information, the subject cannot be easily avoided; statistical training, at least at a basic level, must become an essential ingredient in the balanced education of every citizen.

© New York Times Company, 1982
Reprinted with permission

BEST COPY AVAILABLE

FIGURE 4

Baseball Records and Results

National League

Based on 45 or more times at bats.

	G	AB	R	H	Pct.
J. Thompson, Pitt.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Moreland, Chi.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Bailor, Mets	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Ru. Jones, S.D.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Raines, Mont.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Murphy, Atl.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Lo. Smith, St. L.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Knight, Hou.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Guerrero, L.A.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Wilson, Mets	37	[illegible]	[illegible]	[illegible]	[illegible]

HOME RUNS

Kingman, New York, 13; Murphy, Atlanta, 12; J. Thompson, Pittsburgh, 11; Horner, Atlanta, 10; Hendrick, St. Louis, 9; Guerrero, Los Angeles, 9.

RUNS BATTED IN

Murphy, Atlanta, 37; Kingman, Mets, 35; Moreland, Chicago, 31; J. Thompson, Pittsburgh, 30; B. Diaz, Philadelphia, 27; Horner, Atlanta, 27.

STOLEN BASES

Moreno, Pittsburgh, 25; Dernier, Philadelphia, 21; Lo. Smith, St. Louis, 16; Wilson, Mets, 15; Raines, Montreal, 12.

RUNS

Lo. Smith, St. Louis, 35; Murphy, Atlanta, 34; Horner, Atlanta, 30; Ru. Jones, San Diego, 29; Dawson, Montreal, 27; Kingman, Mets, 27.

HITS

Moreland, Chicago, 55; Wilson, Mets, 50; Lo. Smith, St. Louis, 50; Knight, Houston, 48; Oester, Cincinnati, 47.

DOUBLES

Garner, Houston, 12; Lo. Smith, St. Louis, 11; Knight, Houston, 11; T. Kennedy, San Diego, 11; B. Diaz, Philadelphia, 10; Parker, Pittsburgh, 10; O. Smith, St. Louis, 10; Horner, Atlanta, 10.

TRIPLES

Wilson, Mets, 3; Rose, Philadelphia, 3; Herr, St. Louis, 3; R. Ramirez, Atlanta, 3; Concepcion, Cincinnati, 3; Oester, Cincinnati, 3; Knight, Houston, 3.

PITCHING (5 Decisions)

Forsch, St. Louis, 5-0, 1.000, 3.38; Sutton, Houston, 6-1, .857, 2.40; Ra. Jones, Mets, 5-2, .714, 3.15; Puleo, Mets, 4-2, .667, 3.72; Garber, Atlanta, 4-2, .667, 1.47; Berenyi, Cincinnati, 4-2, .667, 3.38; Welch, Los Angeles, 4-2, .667, 3.31; Rogers, Montreal, 5-3, .625, 1.86.

STRIKEOUTS

Carlton, Philadelphia, 71; Soto, Cincinnati, 66; Rogers, Montreal, 51; Lollar, San Diego, 43; Valenzuela, Los Angeles, 42.

(Last night's games not included.)

Cubs 8, Dodgers 3

THURSDAY NIGHT

CHICAGO — LOS ANGELES

[illegible]

LAST NIGHT'S GAMES

Mets at Houston.
Chicago at San Francisco.
Montreal at Cincinnati
Philadelphia at Atlanta
Pittsburgh at San Diego.
St. Louis at Los Angeles.

EASTERN DIVISION

	W.	L.	Pct.	G.B.
St. Louis	24	15	.615	—
Mets	21	17	.553	2½
Philadelphia	19	18	.514	4
Montreal	18	19	.487	5
Pittsburgh	15	20	.429	7
Chicago	16	22	.421	7½

WESTERN DIVISION

	W.	L.	Pct.	G.B.
Atlanta	25	13	.658	—
San Diego	19	17	.528	5
Los Angeles	19	20	.487	6½
Houston	18	21	.462	7½
San Francisco	17	22	.436	8½
Cincinnati	16	21	.432	8½

(Last night's games not incl.)

TODAY'S GAMES

Mets at Houston (8:35 P.M.) — Falcone (3-1) vs. Ryan (3-6).
Chicago at San Francisco (4:05 P.M.) — Ripley (1-0) vs. Hammaker (2-1).
Montreal at Cincinnati (7:05 P.M.) — Sanderson (3-3) vs. Soto (3-3).
Philadelphia at Atlanta (2:20 P.M.) — Krukow (3-2) vs. Walk (4-3).
Pittsburgh at San Diego (10:05 P.M.) — Sarmiento (1-0) vs. Eichelberger (3-5).
St. Louis at Los Angeles (10:05 P.M.) — Rincon (2-2) vs. Power (0-1).
(All times Eastern daylight.)

TOMORROW'S GAMES

Mets at Houston (n.).
Chicago at San Francisco (2).
Montreal at Cincinnati.
Philadelphia at Atlanta.
Pittsburgh at San Diego.
St. Louis at Los Angeles.

Mets Records

BATTING

[illegible]

Team — AB. 1298, H. 340, HR. 27, RBI. 132 .262

PITCHING

[illegible]

American League

LAST NIGHT'S GAMES

Minnesota at Yankees.
Baltimore at Toronto.
California at Detroit.
Cleveland at Chicago.
Oakland at Boston.
Seattle at Milwaukee.
Texas at Kansas City.

EASTERN DIVISION

	W.	L.	Pct.	G.B.
Detroit	24	12	.667	—
Boston	25	13	.658	—
Milwaukee	19	17	.528	5
Yankees	17	19	.472	7
Baltimore	16	20	.444	8
Toronto	16	21	.432	8½
Cleveland	15	21	.417	9

WESTERN DIVISION

	W.	L.	Pct.	G.B.
Chicago	24	12	.667	—
California	26	14	.650	—
Kansas City	20	16	.556	4
Oakland	20	20	.500	6
Seattle	17	24	.415	9½
Texas	10	23	.303	12½
Minnesota	13	30	.302	14½

(Last night's games not incl.)

TODAY'S GAMES

Minnesota at Yankees (7 P.M.) — Jackson (0-4) vs. Righetti (2-3).
Baltimore at Toronto (1:30 P.M.) — D. Martinez (3-3) vs. Stieb (2-4).
California at Detroit (2:15 P.M.) — Moreno (2-5) vs. Morris (3-3).
Cleveland at Chicago (8:30 P.M.) — Denny (2-4) vs. Trout (3-3).
Oakland at Boston (2:20 P.M.) — McCatty (3-1) vs. Ojeda (1-4).
Seattle at Milwaukee (2:30 P.M.) — Nelson (1-0) vs. Lerch (3-3).
Texas at Kansas City (8:35 P.M.) — Hough (2-3) vs. Gura (4-1).
(All times Eastern daylight.)

Yankees Records

BATTING

[illegible]

PITCHING

[illegible]

Based on 45 or more times at bats.

	G	AB	R	H	Pct.
Bonnell, Tor.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Harrah, Clev.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
McBride, Clev.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
Cooper, Milw.	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]
[illegible]	[illegible]	[illegible]	[illegible]	[illegible]	[illegible]

HOME RUNS

Thornton, Cleveland, 11; [illegible]

RUNS BATTED IN

[illegible]

STOLEN BASES

R. Henderson, Oakland, [illegible]; LeFlore, Chicago, 14; [illegible]

RUNS

R. Henderson, Oakland, [illegible]

HITS

Harrah, Cleveland, [illegible]; Cooper, Milwaukee, [illegible]; LeFlore, Chicago, [illegible]; Garcia, Toronto, [illegible]

DOUBLES

Otis, Kansas City, 14; White, Kansas City, 12; Evans, Boston, 11; [illegible], Kansas City, 11; Cowens, Seattle, 11.

TRIPLES

Herndon, Detroit, 5; [illegible]

PITCHING (5 Decisions)

[illegible]

STRIKEOUTS

F. Bannister, Seattle, [illegible]

(Last night's games not included.)

Tigers 11, A's 3

THURSDAY NIGHT

OAKLAND — DETROIT

[illegible]

BEST COPY AVAILABLE

Red Sox 11, Mariners 2

THURSDAY NIGHT

SEATTLE — BOSTON

[illegible]

Blue Jays 2, Indians 0

THURSDAY NIGHT

TORONTO — CLEVELAND

[illegible]

Orioles 6, Twins 0

THURSDAY NIGHT

MINNESOTA — BALTIMORE

[illegible]

N. B. A.

Eastern Conference

BOSTON VS. PHILADELPHIA

[illegible]

(Philadelphia leads series, 3-2)

*If necessary.
All times Eastern daylight.

Western Conference

SAN ANTONIO VS. LOS ANGELES

[illegible]

(Los Angeles won series, 4-0)

Tennis

AT ROME

ITALIAN OPEN

Quarterfinals

Eliot Teltscher, U.S., d. Eddie Dibbs, U.S., 6-2, 6-3; Andres Gomez, Ecuador, d. Jose Higueras, Spain, 6-2, 3-6, 6-3; Pablo Arraya, Peru, d. Wojtek Fibak, Poland, 3-6, 6-3, 6-2; Mats Wilander, Sweden, d. Tomas Smid, Czechoslovakia, 6-1, 6-1.

AT WEST BERLIN

GERMAN OPEN

Quarterfinals

Bonnie Gadusek, U.S., d. Hana Mandlikova, Czechoslovakia, 7-5, 6-3; Kathy Rinaldi, U.S., d. Mima Jausovec, Yugoslavia, 4-6, 6-3, 7-5; Bettina Bunge, U.S., d. Duk Hee Lee, South Korea, 6-0, 6-1; Sylvia Hanika, West Germany, d. Andrea Leand, U.S., 6-2, 7-5.

FRENCH OPEN
AT PARIS

Seedings for the $489,000 French Open tennis championships which start Monday in Paris:

Men's

1. Jimmy Connors, U.S. 2. Ivan Lendl, Czechoslovakia. 3. Guillermo Vilas, Argentina. 4. Jose-Luis Clerc, Argentina. 5. Vitas Gerulaitis, U.S. 6. Eliot Teltscher, U.S. 7. Peter McNamara, Australia. 8. Yannick Noah, France. 9. Andres Gomez, Ecuador. 10. Balazs Taroczy, Hungary. 11. Brian Gottfried, U.S. 12. Eddie Dibbs, U.S. 13. Jose Higueras, Spain. 14. Steve Denton, U.S. 15. Chip Hooper, U.S. 16. Mel Purcell, U.S.

Women's

1. Chris Evert Lloyd, U.S. 2. Martina Navratilova, U.S. 3. Tracy Austin, U.S. 4. Andrea Jaeger, U.S. 5. Hana Mandlikova, Czechoslovakia. 6. Sylvia Hanika, West Germany. 7. Mima Jausovec, Yugoslavia. 8. Anne Smith, U.S. 9. Bettina Bunge, U.S. 10. Billie Jean King, U.S.

THURSDAY
NATIONAL TENNIS CENTER

LUPUS FOUNDATION TOURNAMENT

Men's Singles

[illegible]

Golf

THURSDAY
AT ATLANTA
ATLANTA CLASSIC

The Leading Scores

[illegible]

College Results

BASEBALL

[illegible]

TOURNAMENTS

N.A.I.A. AREA I

[illegible]

AREA II

[illegible]

AREA III

[illegible]

Soccer

North American Soccer League

THURSDAY NIGHT'S MATCH

Toronto 2, Chicago 0.

Eastern Division

[illegible]

Southern Division

[illegible]

Western Division

[illegible]

TODAY'S MATCHES

San Diego at Montreal (n.).
Tampa Bay at San Jose (n.).
Toronto at Chicago.
Tulsa at Jacksonville (n.).

Six points awarded for a regulation or overtime victory. Four points for a shootout victory. One bonus point for every goal scored with a maximum of three per game. No bonus point is awarded for overtime or shootout goals.

M.I.S.L.

FINAL

NEW YORK VS. ST. LOUIS

May 14 — St. Louis 3, New York 2 (overtime).
May 16 — New York 3, St. Louis 2.
Last night — at St. Louis.
May 23 — at St. Louis, 8:30 P.M.
*May 26 — at New York, 8 P.M.
(Series tied, 1-1)

*If necessary.
All times Eastern daylight.

A.S.L.

LAST NIGHT'S MATCH

Nashville at Georgia.

[illegible]

(Last night's match not included.)

TONIGHT'S MATCH

Oklahoma City at Carolina.

Five points awarded for a victory, two for a tie, and one bonus point for each goal scored up to a maximum of three.

Lacrosse

N.C.A.A. MEN'S CHAMPIONSHIP

SEMIFINALS

Today

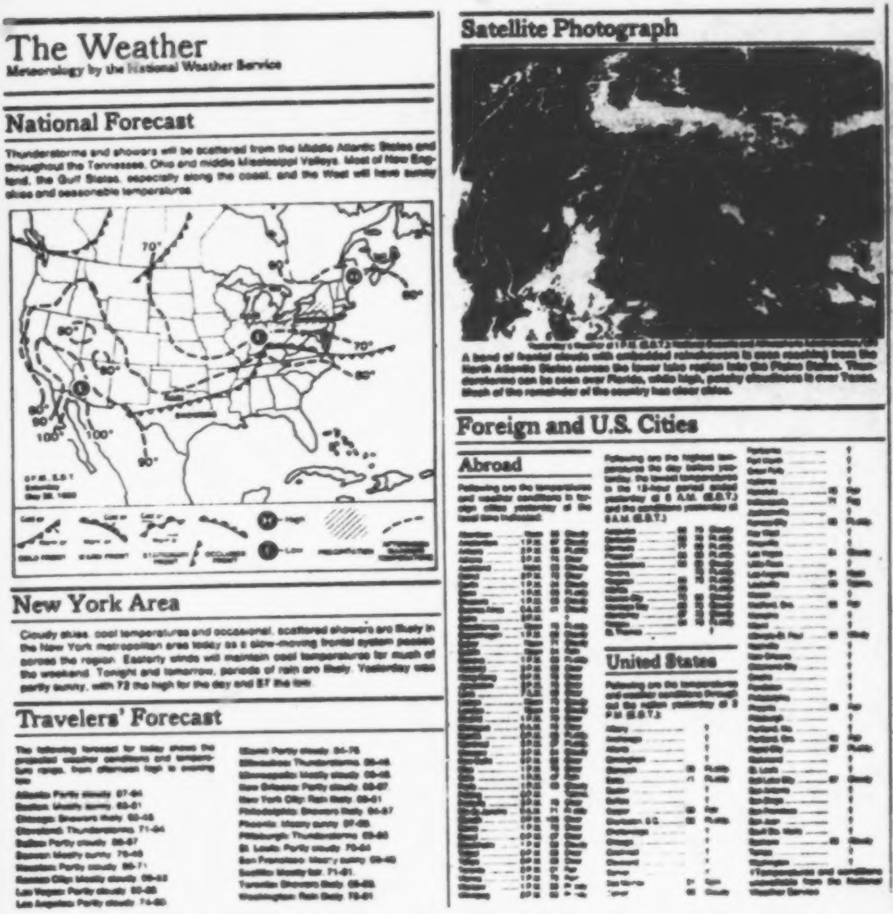
North Carolina at Cornell.
Virginia at Johns Hopkins.

Final

May 29 — at Charlottesville, Va.

BEST COPY AVAILABLE

FIGURE 5


70°
80°
90°
100°
High
Low
Cold front
Warm front
Stationary front
Occluded front
Precipitation

The Weather

Meteorology by the National Weather Service

National Forecast

Thunderstorms and showers will be scattered from the Middle Atlantic States and throughout the Tennessee, Ohio and middle Mississippi Valleys. Most of New England, the Gulf States, especially along the coast, and the West will have sunny skies and seasonable temperatures.

New York Area

Cloudy skies, cool temperatures and occasional, scattered showers are likely in the New York metropolitan area today as a slow-moving frontal system passes across the region. Easterly winds will maintain cool temperatures for much of the weekend. Tonight and tomorrow, periods of rain are likely. Yesterday was partly sunny, with 72 the high for the day and 57 the low.

Travelers' Forecast

The following forecast for today shows the projected weather conditions and temperature range, from afternoon high to evening low.

Atlanta: Partly cloudy. 87-64
Boston: Mostly sunny. 63-51
Chicago: Showers likely. 65-48
Cleveland: Thunderstorms. 71-64
Dallas: Partly cloudy. 88-67
Denver: Mostly sunny. 75-48
Houston: Partly cloudy. 86-71
Kansas City: Mostly cloudy. 68-53
Las Vegas: Partly cloudy. 90-65
Los Angeles: Partly cloudy. 74-60
Miami: Partly cloudy. 84-76
Milwaukee: Thunderstorms. 66-48
Minneapolis: Mostly cloudy. 68-48
New Orleans: Partly cloudy. 88-67
New York City: Rain likely. 66-61
Philadelphia: Showers likely. 64-57
Phoenix: Mostly sunny. 97-65
Pittsburgh: Thunderstorms. 69-60
St. Louis: Partly cloudy. 75-64
San Francisco: Mostly sunny. 69-49
Seattle: Mostly fair. 71-51.
Toronto: Showers likely. 66-53
Washington: Rain likely. 70-61

Satellite Photograph

A band of frontal clouds with embedded rainshowers is seen stretching from the North Atlantic States across the lower lake region into the Plains States. Thunderstorms can be seen over Florida, while high, patchy cloudiness is over Texas. Much of the remainder of the country has clear skies.

Foreign and U.S. Cities

Abroad

Following are the temperatures and weather conditions in foreign cities yesterday at the local time indicated.

Following are the highest temperatures the day before yesterday, the lowest temperatures in the 12-hour period ended yesterday at 8 A.M. (E.D.T.) and the conditions yesterday at 8 A.M. (E.D.T.)

United States

Following are the temperatures and weather conditions throughout the nation yesterday at 8 P.M. (E.D.T.)

†Temperatures and conditions unavailable from the National Weather Service

© New York Times Company, 1982
Reprinted with permission

A CORE OF STATISTICAL TRAINING FOR THE LAYMAN

What are the basic statistical concepts required for an intelligent reading of The New York Times of 22 May 1982? Can we extract from this single newspaper issue, and from a composite of other ordinary information material, a core of statistical topics which would assist the layman to a deeper understanding of our world?

The first point that must be made is that while the terminology may vary in different areas such as finance, economics, social science, politics, and natural science, the statistical principles which are applied to numerical data in each of these fields are identical. They are, in effect, largely context-free. The material we have examined above requires some grounding in the following topics:

1. Descriptive statistics: numerical data and ranges of variation
2. Graphical statistics: graphical representation of data (including time series).
3. Economic statistics: economic indices and their rates of change.
4. Demographic and social statistics.
5. Polls, sample surveys and voting formulae.
6. Concepts of statistical inference.
7. Errors and fallacies in statistical arguments.

For the sake of completeness, one might add for good measure:

8. Statistical calculations: mean, variance and the concept of a distribution.
9. Elements of correlation and regression.
10. Probability models and how to fit them to statistical data.
11. Basic concepts of insurance statistics.

This will undoubtedly seem a very ambitious program, but it is one which is essential to a proper understanding of the information we encounter, and to the conduct of our daily life. If we intend to prepare our students to grapple with the problems of society, we have little choice but to include such a core of statistical training in their schooling and college education. Because of its dependence on numeracy, such a program must inevitably be carried out in conjunction with the students' development of arithmetic and computational skills.

How can this best be done in our schools and colleges? It would be tempting to insist that we set up formal courses in statistics to cover the above topics, as part of every school and university curriculum; we could then demand that every student attend them. Such an attempt would frighten many, particularly those with modest numerate skills, and might well prove counter-productive in the long run. Possibly a lower-keyed approach on multiple fronts would prove more successful. Mathematically oriented courses in statistics should certainly be provided for those who can benefit from them. But one must simultaneously cater for students of modest numeracy by providing them with a more statistical orientation in business, biology, economics, geography, history, politics and social science courses, as well as all the other fields of study offered in our educational system. In doing so, one would discuss most of the 11 topics previously outlined.

Since statistics enters into the subject matter of most fields in which quantitative data arise, it is important that students should be able to understand its uses, and learn to handle simple statistical problems with some facility. Several statistical text books exist at a variety of levels, both school and college, which endeavour to achieve this aim, and which teachers have found useful as guides. But they are mostly published in the developed countries of the world; it will be one of our challenges to explore whether they can be adapted to the requirements of developing nations, or whether completely new texts will need to be written for these.

Text books may be important in our drive for statistical literacy but skilled teachers will prove crucial to its success. The shortage of such teachers is one of the difficulties experienced in the developed countries over the past two decades, when the role of statistics in education has begun to achieve some recognition. There is no surplus of school and college teachers competent to teach statistics, or to inject a statistical orientation into courses in other fields of study. Not all teachers of mathematics are fully familiar with basic statistical concepts; biologists, geographers and social scientists, among others, are often not sufficiently confident of their statistical skills to discuss statistics problems as they arise in their courses.

Hence, to succeed in educating our youth statistically, we must begin by educating our teachers. In this campaign, the production of suitable text books, the publication of such magazines as The American Statistician, and Teaching Statistics, the organization of Conferences such as ICOTS, and the recommendations of the ISI Committee on Statistical Education and UNESCO all have a valuable role to play. But the most important factor of all will be the teachers' recognition of the need for statistical literacy, and their personal willingness to retrain themselves to respond to the problem. Such retraining may require them to attend appropriate university statistics courses; they may need to convince their Education Departments that statistical material should be included in various student programs of study, and plan to modify syllabuses gradually so that statistics becomes an integral part of other subject courses. If the ISI Committee on Statistical Education can assist such progress in any way, please ask for its help.

We are envisaging an enterprise of the utmost importance: a core training program for world-wide statistical literacy. This is an endeavour linked with the spread of numeracy, which could well take several generations to bear fruit. To carry it out, we must train our teachers first, so that they can in their turn train our youth at both the school and college levels. Such training will require detailed and careful organisation, with different emphases in different countries. In bringing together teachers of statistics of all nations, ICOTS has made a contribution to this long-term education process.

CAN BASIC STATISTICAL TRAINING BE CULTURE-INVARIANT?

We have put forward the suggestion that, for the majority of students, it may be more appropriate to introduce statistical concepts and techniques into the teaching of other subjects, rather than set up a universal basic course in statistics. We must now face the question of whether such training can be culture-invariant. Do the citizens of developed countries require the same knowledge of statistics as those of developing nations? Will the basic topics of statistical training vary depending on whether we teach Nigerians or Japanese, Russians or Upper Voltans? Clearly there are many responses to this question; it can be answered partially by stressing the universality of statistical theory and methods. But another partial answer will point to the differing needs of individual countries, as well as to their educational and economic resources.

If we concentrate purely on the concepts and methods of statistics, as embodied in the 11 topics of Section 3, then we can reply very briefly that basic statistics is independent of culture. Means and variances, significance levels, and the fitting of regression lines are identical for all countries: statistical procedures have the same culture-free validity as the other methods of science. But to the extent that statistical data arise in an economic or social context, rather than a scientific one, then they may well vary from country to country. It is relevant here, to point out that neither the Sunday Times of Nigeria of 6 June 1982, nor the Overseas Hindustani Times of 31 December 1981, published as large a proportion of statistical material as The New York Times. In both of them, approximately 10% of the material involved elementary descriptive statistics. It would appear that in 1982, technically advanced countries were more statistically oriented than agricultural ones, even those with vigorous nascent industries.

Even in developed nations, examples and illustrations of statistical principles may differ substantially. While it might be sensible to discuss variation and trends in economic time series by reference to the stock market in the USA or Britain, such concepts would perhaps be more suitably discussed in terms of annual oil or steel production in the USSR, where the stock market does not exist. In developing nations, statistical discussion must center on topics of vital concern for the particular country. In Egypt, for example, one might select the statistics of fertility and mortality, data on the cotton crop, or the growth of oil production, while in the Dominican Republic the sugar crop will clearly provide statistical data of direct interest. In short, the discussion and analysis of statistical data must reflect local conditions and interests, if statistics is to become an integral part of every citizen's education. Statistics text books and courses in different countries will show signs of such cultural differences; even if a standard English text is adapted for use in a developing country, its examples may require modification to make the contents more interesting to its new students.

A more fundamental point concerns the differences in economic and educational resources in different countries, and the consequences of these for a statistical training program. There is an order of human priorities which must be kept clear in our minds: it makes little sense to press for statistical sophistication in a country where unemployment is persistently high, where literacy is only partial, and where disease is endemic. Calcutta is the home of the famous Indian Statistical Institute; yet I would hardly rate the statistical education of Calcutta's citizens as a first priority. Like all other skills, statistical expertise subsumes health, an acceptable standard of living, and a reasonable level of literacy and numeracy. We must not, like Dr. Pangloss in Voltaire's "Candide" get carried away with the beauties of philosophy (or statistics) in the midst of famine, carnage and human injustice. In this sense, statistical training cannot be culture-invariant.

Most European countries are considered to be developed. They have well established school and university systems, with a trained body of teachers in mathematics and the natural and social sciences, capable of handling the basic statistical training of students which we have proposed, either immediately, or within a relatively short period. It is no accident that Teaching Statistics was published in Britain, or became immediately popular in Germany, Scandinavia, and other European countries. The ground there is ripe for statistical literacy.

The same is not true, for example, of India or South America. Here, with some notable exceptions, it would be difficult to introduce statistics in schools at this stage; it is only in some of the universities that progress in the spread of statistical education can be made. But if progress can be made in the universities in this generation, and some basic statistical training can become part of every college student's education, then there is hope that such education will filter downward in the next generation, as the standards of living and education rise. Every country must decide for itself how best statistical literacy can be increased, and which educational sector is most likely to be responsive to the growth of statistical education. And in every country, school and university teachers interested in statistics must try to convince their governments of the value of laymen's understanding of the subject.

This postulates that every government is keen to rule over an intelligent and well informed people. One may cynically observe that this is not always so, either in the democratic or in the totalitarian countries. In this sense, the drive for statistical literacy may be only one of the facets of the human struggle for social justice, and for the fulfilment of every living person's potential. These are sentiments nobly enshrined in the United Nations Universal Declaration of Human Rights, but more often honoured in the breach than in the observance.

To summarize our views, one might say that while basic statistics is culture-invariant, the case studies by which statistics will be illustrated may be culturally oriented. It is up to us, as concerned statisticians in every country of the world, to devise the best methods of imparting statistical literacy to our youth, to produce the best texts that we can with relevant local examples, and to incorporate in these the core of statistical principles which we have previously outlined. The ISI Education Committee is vitally interested in this campaign for statistical literacy, and will give the fullest possible assistance in its pursuit.

CONCLUDING REMARKS

Statistics is relevant: in a world where science and technology are rapidly assuming a dominant role, it is impossible to understand the daily newspaper without a smattering of statistical knowledge. To achieve this, we have suggested a basic core of statistical literacy, and a program of teachers' training, followed by the widest possible statistical education of students in schools and universities. We have tried to set this campaign for statistical literacy within the broader context of the struggle for human rights.

And now, we face a further revolution, a revolution not only in our technology, but also in our methods of supplying, storing and using information. While the electronic computer has been with us for over thirty years, it is only recently that mini-computers of great power have entered into our homes in the developed countries. The basic Sinclair computer, which hooks up to a TV set, costs only $150; others like the Apple III, or the TRS 80, or the Commodore PET cost at most a few thousand dollars, and are rapidly decreasing in price.

These mini-computers handle numerical and statistical data with great facility, calculate means and variances, plot histograms, detect trends in time series, carry out standard significance tests, and generally do all that a sophisticated statistician might require of them. The layman may not understand the finer details of the procedures which the computers are programmed to follow, but he can certainly understand and use the results. It is therefore important that the layman be made aware, not only that there are statistical fallacies, but also bad statistical programs. To statistical literacy, we may in future have to add programming literacy.

These two may gradually become confounded, but for the moment let us meet one problem at a time: we should begin with a campaign for numerical and statistical literacy, to equal in this 20th century that for literacy itself which occurred during the 19th century. When we have made some headway with it, it will be time enough to turn to computer and programming literacy. Life constantly renews itself, and we can rely on it to provide us with new challenges tomorrow, as soon as we have resolved the problems which we face today. Let us concentrate in the 1980s on statistical literacy, and hope that within our own lifetimes we shall see the results of our efforts, and the consequent spread of such literacy throughout the world.

TOPIC 1: TEACHING STATISTICS IN SCHOOLS
(6-11 YEARS AGE GROUP)

ORGANISER: PROFESSOR K MATUSITA

STATISTICS IN THE CURRICULUM FOR EVERYBODY - HOW YOUNG CHILDREN AND HOW THEIR TEACHERS REACT

Tamás Varga, National Institute of Education, Budapest

STATISTICS in the title means more than simple descriptive statistics such as representing children born in a given month or families with a dog or a cat by piled up matchboxes and the like. But it means less than any serious inferential statistics, if only

because

CURRICULUM FOR EVERYBODY in the title is meant to be the one implement in the first four grades of every Hungarian school six to ten year olds, where statistics, in fact, appears at every grade level along with combinatorics and probability, as also in the subsequent fifth to eighth school year, where the implementation is still under way, and also

because

YOUNG CHILDREN are meant to be the said six to ten year olds though most classroom examples will be related to a particular class of third graders in Offenbach,* West Germany, nine year olds on average, with whom I decided to gather some fresh experiences while preparing this talk. As for the TEACHERS, I have in mind mainly those I regularly meet in Hungary during in-service training courses. But the bulk of what I am going to say is about children.

Back to the first word in the title, the reason for going beyond simple descriptive statistics as early as in the first few primary grades is that young children seem to be more responsive to essential, deep questions than to occasional, superficial ones. They are eager to study random phenomena, which they sense to be worthy of their attention. Questions of birth and life and death, accidents and disasters just as well as funny coincidences rouse their interest. So do random phenomena involving actual experiences in their world of plays and games. Their world also includes tales: stories and beings of their imagination, which can be modelled in their games. One more facet of their interest is that they like personal involvement. Pet graphs appeal to them if their own pets are included. In their experiences about random phenomena their involvement is greatly increased by first predicting the outcome then performing the experiment, finally comparing the result with their prediction. In some cases a further step is an attempt to explain the pattern of results by some mini-theory. This carries them further toward probability. During their games they are motivated to devise strategies which increase their chances, and then they get some feeling for game theory. The mini-theories demand combinatorial considerations. Yet since the order is

---

* I owe big thanks to the classroom teacher, Mrs. Sommer, for allowing me to take over her class of 14 girls and 11 boys, coming from seven different countries, during eleven times 45 minutes and sometimes longer.

hypothesis ⟶ experience ⟶ theory,

I feel it appropriate to subsume such learning sequences under the term statistics.

The present talk is based on extensive experience, but no controlled experiment, no statistical data, no research like that done by Piaget, Estes or Fischbein.

It is just my private opinion that children usually behave as if they have read their books and knew their roles related to their ages. This is particularly so in the beginning of a sequence of lessons. If, however, the lesson inspire them - and I hope it is really inspiration rather than indoctrination - then they may forget their age and role, and behave at a more mature level. This should, of course, be tested, and possibly in longitudinal research, in order to prove how lasting the effects are.

Now I start to report on certain segments of actual lessons, intermittently adding my remarks.

TWO LANGUAGES FOR PREDICTING LEVELS OF CERTAINTY

Teacher: 'We are going to study probability. It is about things you cannot be sure of. They are in between what is certain and what is impossible. For instance tomorrow will certainly be a Wednesday, which I can show this way (shows a stretched palm with fingers upward). It is impossible that tomorrow it will be Monday (stretched hand, fingers downward). But you cannot be sure whether it will rain or not. You can at most be pretty sure, but not absolutely certain. (Stretched hand, fingers slantly upward). Or you may feel, it is rather unlikely to have rain tomorrow. (Slantly downward). Or maybe you feel so: it is just as likely that it will rain as that it will not. How would you show that? (Some children stretch their fingers forward, others wave them up and down). Discussion follows about the possible interpretations of these two gestures. 'If you wave your hand, then I will think that you change your mind in every moment'.

Examples come next, imaginary and actual events where children express their opinions or predictions by the hand language. Random devices are used: coloured beads to choose of, blue-red counters to throw, etc. Later - maybe in a few days - children learn a written (or rather drawn) language to express their predictions. They do it by stick pictures like these: (Fig.1).

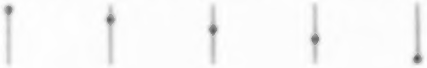


FIG. 1

or these (Fig.2)

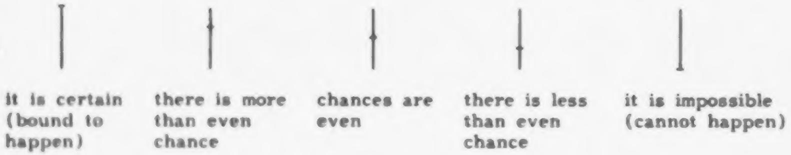

it is certain (bound to happen)
there is more than even chance
chances are even
there is less than even chance
it is impossible (cannot happen)

FIG. 2

Two important cases are near the two extremes: (Fig.3)

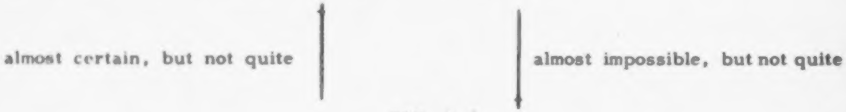

almost certain, but not quite
almost impossible, but not quite

FIG. 3

WHY THESE LANGUAGES?

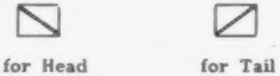
Both the gestures and the stick pictures are vehicles of non-verbal communication. The gesture language enables instantaneous, simultaneous, and - last but not least - silent feedback in a class of children. The stick pictures are initially a rough representation of children's probabilistic intuitions. The sticks can be drawn of any height; it is the relative position of the marking that counts.

The use of square grid (graph paper) allows a more refined quantification, and paves the way to representing the distribution of random variables.

Both languages suggest the idea of probability as a quantity between two extremes, while avoiding the sophisticated idea of numbers between 0 and 1 - and at the beginning totally avoiding the use of numbers.

CHILDREN SIMULATE COINS

Every pupil receives a piece of graph paper with a rectangular frame around 8 times 16 square boxes. Their task is to figure out and record the behaviour of a randomly tossed coin. One way of doing this is to fill each of the eight rows of 16 squares from left to right one after the other by two different slant lines like these:



The choice of these two symbols results from an attempt to minimize both the time of recording and the difference between the effort to record Head and Tail. The results are found in Table 1 arranged in decreasing number of runs of length 1. (A boy was absent). For easy comparison, expectations in 128 actual tosses are given in the last columns.

TABLE 1

		1	2	3	4	5	6	7	8	9	10	11	12	
		B	B	B	B	B	B	G	G	B	G	G	B	E.
Runs of length	1	101	73	69	66	60	58	55	54	47	42	41	41	32
	2	12	18	7	6	28	10	35	23	34	9	28	17	16
	3	1	5	2	14	4	4	1	8	3	8	9	9	8
	>3	0	1	7	2	0	7	0	1	1	8	1	6	8
Runs		114	97	85	88	92	79	91	86	85	67	79	73	64
Tosses in runs of length	1	101	73	69	66	60	58	55	54	47	42	41	41	32
	2	24	36	14	12	56	20	70	46	68	18	56	34	32
	3	3	15	6	42	12	12	3	24	9	24	27	27	24
	>3	0	4	39	10	0	38	0	4	4	44	4	26	48
Tosses		128	128	128	128	128	128	128	128	128	128	128	128	128
Heads		64	51	63	67	65	59	61	62	62	72	68	65	64
Tails		64	77	65	61	63	69	67	66	66	56	60	63	64
Difference		0	26	2	6	2	10	6	4	4	16	8	2	≈9

		13	14	15	16	17	18	19	20	21	22	23	24	
		B	G	G	B	G	G	G	G	G	G	G	G	E.
Runs of length	1	40	39	39	37	34	34	31	29	27	22	17	11	32
	2	16	35	28	6	23	16	30	19	14	7	17	8	16
	3	9	5	11	2	10	15	5	6	13	9	14	13	8
	>3	7	1	0	12	4	4	5	10	8	13	8	14	8
Runs		72	80	78	57	71	69	71	64	62	51	56	46	64
Tosses in runs of length	1	40	39	39	37	34	34	31	29	27	22	17	11	32
	2	32	70	56	12	46	32	60	38	28	14	34	16	32
	3	27	15	33	6	30	45	15	18	39	27	42	39	24
	>3	29	4	0	73	18	17	22	43	34	65	35	62	48
Tosses		128	128	128	128	128	128	128	128	128	128	128	128	128
Heads		60	65	67	56	66	65	69	65	68	63	64	62	64
Tails		68	63	61	72	62	63	59	63	60	65	64	66	64
Difference		8	2	6	16	4	2	10	2	8	2	0	4	≈9

For those who have ever administrated a test like this to children - or, for that matter, to uninitiated adults - it will not be surprising that most of them believe shorter runs to be more frequent than can be expected in actual tosses. The different behaviour of boys (B) and girls (G) is in this small sample quite striking - girls seem to believe in more constancy on the part of a coin than boys, though most of them avoid very long runs, too.

COINS - FAIR OR UNFAIR?

The teacher prepared some loaded coins and dice, along with practically fair ones. The loaded coins, and some unloaded ones too, were rather thick, with real coins on one side, cardboard on the other, the whole of them wrapped, and marked by a 0 and a 1. The whole class assisted while a coin produced these outcomes:

00100 10100 00010 01110 11000 11101

So far there were 17 zeroes (Heads) and only 13 ones (Tails). Yet, in the last five tosses all but one was a Tail. The majority were Tails in another round of five. Is the coin loaded or fair? Or should we go on with the tosses, before taking position?

Children voted in favour of going ahead, doing 30 more tosses. Before continuing, the teacher asked them about the numbers of Tails in 60 tosses they felt would be still acceptable on the part of a fair coin. Would the coin be still rather suspicious if in 60 tosses there were only

20, 22, 24, 25, 26 tails?

Up to 25 all of them said they could not believe that the coin was fair. At 26 some of them said it could be. At this point the teacher tried to make the meaning of the vote more precise:

... 22 23 24 25	26 27 28 ...
In all these cases the coin would be maintained by them to be suspicious.	In such cases they could still accept it as being fair.

Then they went on with the numbers. At 27 the majority voted that from this onward numbers were already acceptable. Some votes were cast to 28 and 29. (There was no mention of numbers greater than 30). Most children presumably had just a very vague idea of what it was all about, but the teacher felt that these vague ideas moved them in the proper direction.

After 30 more tosses the end result was: 24 Tails. The coin was declared by all of them to be suspicious. Cut it! Cut it! - they yelled. The teacher cut the coin, and pulled out two big Hungarian coins from one side.

He was somewhat deceived - not with the children, but with the coin. He expected a more clear cut majority of Heads. At home he found a relative frequency very near to 2/3 in several hundred tosses. There he kept the conditions strictly constant by dropping the coin on the floor from a distance of about 120 centimeters. Was the floor in the school much different? Or did the fact that he now tossed it change the chances? Maybe it did. But even under precisely the same conditions the difference between 24 and 20, only slightly beyond σ, was hardly enough reason for any great surprise, let alone for modifying his original assumption about the approximate probability. Most children expected still more conformity on the part of the coin in order to accept it to be fair - well within one σ without, of course, knowing anything of the binomial distribution and its S.D. At their level of knowledge they still behaved more intelligently, their probabilistic intuitions worked better, than that of their teacher. Sixty tosses are clearly too few to distinguish between two conflicting assumptions such as 1:2 or 1:1 odds - rejecting any of the two in favour of the other - at an acceptable level of significance.

Soon after this the falsity of a strongly loaded big white cardboard die was quickly detected: after a majority of sixes they cut it and found a lot of plasticine below the one-dotted face.

HOW DIFFICULT IS IT TO THROW A SIX?

The teacher started with the two Freudenthalian questions:

1. Which is easier to throw, a one or a six?
2. To whom is it easier to throw a six, you or me?

Most children were acquainted with a game where players could start moving their pieces only after they threw a six. The memory of their impatience may be one of the reasons why initially many of them believed it was more difficult to throw a six, than - say - a one. But the opinions were not unanimous. Nor were they about the second question. These children were not totally uninitiated.

I will try. You will count the number of throws until the six appears. Who of you believes that this will already happen the first time?
He writes the number 1 on the board, but no name goes below it. Nor below the 2. Some believe that the six will first appear at the third throw the teacher writes their names under the number 3. Most names get below the 4 and the 5, some below the 6, one below the "10 or more".

Let us see! Incidentally the first throw is a six. Tally mark below the 1!. Let us start again: one, two, three. The third is a six. Those who predicted are the champions, they get gold medals. Tally below the 3. So they go on together - some get silver, some bronze medals - then they continue the game in groups of two or three: one of them throws, another writes the tallies after the numbers (this time in a column, on the left side of their worksheets, from 1 to 20 or more) a third may just look and check if everything is OK. After ten minutes: dice back, outcomes are pooled. One in each group shows with his fingers, how many tallies are after the 1's; these numbers are added. Then after the 2's etc. Unbelievable! The 1 got the most tallies, then the 2.

There are good chances of this to be the case anyway - if a few hundred tallies are pooled - before oscillations prevail.

Through this activity children get a feeling of
a) the interplay between law and casualness,
b) a law - geometrical distribution, where mode, median and mean, here 1,4 and 6, are clearly distinct.

This will help them to grasp the meaning of each of these, now confusedly present in their intuitions. They cast their votes to numbers around the median (when does it start to be more likely to have thrown a six than not to have yet thrown it?) or the mean (theoretically: the expected number of throws until a six appears), without realizing, that their winning chances can be maximized by asking another question: the one about the most frequent (theoretically: the most likely) number of throws until the six appears.

FIRST DIGITS

Which is the most frequent letter in German?
Quite a few children voted for E, and the teacher showed on the top of a paper roll he started to unroll that it was in fact the E with 173 appearances in 1,000 letters. The top of the roll was like this:

E	173
N	104
R	81
I	75
S	64

Is the E the most frequent letter at the beginning of words as well?
Some said yes, some voted for D or other letters. Let us see! Open your reading books at random, read a passage chosen at random, do the tallying after the letters you have on your worksheets, alphabetically arranged.

After a few minutes the teacher asked them to dictate - the numbers appeared on the board - and then add the frequencies first of the E's, then of the D's as initial letters. The D's got a clear majority.

Maybe there is some other letter more frequent than the E as initial letter?
They tried the S and the T, also the W and a few others. Then the teacher changed the topic and asked:

What are to numbers as letters are to words?
A lot of them knew and shouted: digits! Which is the most frequent digit?
The O, said a boy, and it was a reasonable answer. They discussed that O's very often appear at the end of numbers, though in prices 9's compete with them, or they even prevail.

Which is the most frequent first digit?

The digit 1 was pronounced. A girl started to list products, mainly foods, the unit price of which was between DM 1 and DM 2. Others were not convinced, and started to list further foods, with prices higher than DM 2.

This was the end of the lesson, and the teacher had only time to let children choose from a thick bundle of newspaper clippings mainly advertisements of foods, furniture, trips, sport articles, also radio and TV programmes etc.. He invited them to examine first digits. Some children undertook a different task: to open a telephone directory and do the tallying about the first digits of the house numbers. It was made clear that they were expected to prepare tally lists of this format:

1	
2	1
3	11
.	
.	
.	
9	

Next day most of them came to school with the finding that 1 - and possibly 2 - were more frequent than any other first digit, in quite different situations.

During the lesson some more situations were added. One of them was within mathematics: the multiplication table from 1 x 1 to 10 x 10, previously written on the board. Others were taken from life, e.g. the dates of birth of children, and possible dates of birth in general. Let us forget of the years; their first digits are always 1 (for the time being, and not much back in time). In four of the twelve months the first digit is 1 - the other eight digits have an equal share. In the possible days of the month 1 and 2 have an equal share with eleven possibilities each, 3 has only three possibilities of being a first digit, the other digits have only one.

A similar argument explained the case of TV schedules.

How about house numbers? Is there an upper limit? This was a more complicated situation, it needed some idealisation. Imagine a town where all the streets are of different length. In one of them there is only one house, its number is 1. Call it street one. In another street, call it street two, there are two houses, number 1 and number 2. And so on. Are there in any of these streets an equal number of houses with each of the nine possible first digits?

Yes, in street 9 each of the digits 1 to 9 appears as a "first" digit exactly once. It was harder to detect that the same was true about street 99. It was not important to go on to 999 etc.

Is there any street, where the champion on the first place is the 9? Or the 8? Or, for that matter, the 2?

They realized that none of these could happen. Even the 2 could only climb up to sharing the first prize with the 1, not to be superior.

The argument was not pursued further, it remained at this qualitative level. The invariance under a transformation - such as the areas of countries in square kilometres and in square miles - was kept for a higher level.

FROM PERCEPTION TESTING TO HYPOTHESIS TESTING

Could any of you distinguish these sorts of chewing gum if blindfolded? Some said they could. One of them chose two gums, one wrapped blue, the other green. She was granted a short learning period, half a minute or so. This the teacher felt necessary, to avoid much change in the probability of hitting upon the correct answer in the course of trials. (Immediate feedback information about the hits had to be given to the subjects, if only because of the others who cheered at hits and sighed at failures). She succeeded in five trials in a row, and this the others accepted as a perfect proof of her ability.

A second volunteer, a boy, succeeded in three trials out of five, but failed in two. Some children wanted to be nice and tended to accept his performance saying that he still more often succeeded than not. The teacher warned them: How can you tell that the 3 to 2 result was more than a little luck? But he understood that tests administrated this way would not be highly motivating. He postponed this kind of activity until a later session. He prepared various test materials where the perceptive difference could gradually be decreased. This would give subjects an initial experience of success in cases where the differences were considerable, and a feedback about their perceptive ability levels which, of course, could not be unlimited. The following types of tests were devised, not all of them used so far:

Length estimation:
Rectangles with slightly different sizes were displayed. The subjects had to tell whether or not the vertical sides were the longer ones (whether or not the rectangles were turned by a right angle while the subject could not see them).

Number estimation:
Cuisenaire rods of the same size (and hence of colour, too: the red ones were used) on two trays, were shown a few more on one of the trays. Subjects had to tell whether it was the one to the right or the left. They had only one second to inspect the trays, and the number of rods was at least 15,

to avoid counting. The rods were not arranged in any regular pattern, and their positions were also changed, not only their numbers.

Weight estimation:

Rods were used again, but in closed boxes, and children had to distinguish which was heavier by lifting them.

Time pitch and taste estimations were also planned; the latter with a small quantity of salt - or of sugar - in one cup, to be distinguished from the other, just water.

From the point of view of teaching statistics, the reactions of children are more interesting than technical details. Batches of five trials were usually administered. Children felt that the following qualification is fair:

No failure in five trials : excellent
One failure in five trials : acceptable

How many in ten trials do you think would be acceptabe? the teacher asked. The answer was again almost immediate: four.
The teacher tried to make clear that "acceptable" meant hardly by pure chance, and asked them, if they felt that two or less failures in ten trials could be the result of guesswork just as easily as one or none in five. They did.
He decided to postpone confronting them with contrary experimental evidence at a later date.
This was done by letting them shake boxes containing five, ten, or twenty two-coloured counters. (They were very small but relatively thick, to avoid overlapping). Without any calculations these experiments convinced children that their reasoning through proportionality - which they never learned but intuitively felt to be correct - was hardly applicable in situations like these.

Some tentative conclusions:

1. Statistics as a school subject is quite suitable in the primary school.
2. Children need not know much of arithmetic before they are initiated into statistical ideas.
3. Proportionality is not an indispensable prerequisite to understanding statistical - or probabilistic - ideas. It can even be a hindrance, and not only after being conditioned to look at proportionality as a key to most mathematical questions.
4. Learning statistics enhances children's arithmetic understanding and skills.
5. Learning statistics enhances, in particular, their ability of applying arithmetic to life situations.
6. Learning statistics enhances a number of other abilities related to mathematics and its applications, not just arithmetic ones.

TEACHING OF STATISTICS IN ELEMENTARY SCHOOLS IN JAPAN

T Yanagawa, Kyushu University

INTRODUCTION

Mathematics, science and social studies are the three major subjects which have the responsibility for teaching statistics in elementary schools in Japan. The basic and systematic teaching of statistics occurs in the mathematics classes. Students are exposed to various kinds of statistical charts, graphs and tables in science and social studies. They are supposed to study there how to use charts, graphs and tables for communicating information clearly and effectively. although these three subjects are taught by one teacher, textbooks are edited separately and independently. The statistics taught in these subjects and their interrelation are mainly results of a history of trial and error. The present system has promoted the teaching program of traditional topics such as charts, tables and graphs. But, at the same time, it has suppressed the introduction of new ideas into the program; in fact, statistics taught in elementary schools in Japan in the past three decades has changed only a little.

Japan has changed drastically during these three decades. Today we live in a highly developed technological society. Students are exposed to a great deal of numerical information in their daily lives. The importance of probabilistic and statistical ideas has thus increased considerably. It seems unreasonable that the essence of statistical education available in elementary schools today is almost the same as that of three decades ago.

The influence of Japanese statisticians on the teaching of statistics has been insignificant; we still have no research and teaching institute such as a Department of Statistics in Japanese Universities. The following are examples of our efforts and indications of the improvement in the environment and the quality of statistical teaching in Japan:

(i) The Japan Statistical Society established a committee of statistical education about 10 years ago and has delivered a lecture/consultation every year to school teachers, at the site of the annual meeting, hoping to improve the environment and the quality of statistical education.

(ii) For the first time in our history, the name of a statistician can be found among the 5-10 authors of each of the latest high school mathematics textbooks. We hope this will influence the teaching of statistics in elementary schools.

(iii) Teachers colleges have begun to employ statistical scholars in recent years. It is estimated that about 17% of teachers colleges have staff members who have published at least one theoretical or applied paper in statistics.

The purpose of the present paper is to describe how, what and when statistics is taught in elementary schools in Japan and to discuss its advantages and disadvantages with several suggestions for its improvement. In section 2 we will show the objectives and plans for the teaching of statistics in elementary schools established in the guideline of mathematics issued by the Ministry of Education. The guideline is detailed and all textbooks must pass its examination before they are published. In section 3, looking into mathematics textbooks, we will show that textbooks of mathematics are specialized for the systematic and efficient teaching of mathematics. Traditional topics in statistics such as charts, graphs and tables, and their prototypes, are efficiently employed for introducing and developing important ideas in mathematics. We will show in the section how beautifully such topics may be taught in mathematics. It is important to give students experience in collecting and analyzing data. However, the collection of data is neither encouraged in the guideline nor in the textbooks. We need to establish the program regarding the collection of data, coordinating with all subjects taught in elementary schools. Further the other important topics in statistics which cannot be incorporated in the systematic teaching of mathematics such as the role of the mean for central tendency, the dispersion, the idea of sampling and that of probability are poorly handled in most mathematics textbooks. We will look into these problems in section 4 with several suggestions for its improvement.

THE GOALS OF THE GUIDELINE

The objectives and plans for teaching statistics in elementary schools in Japan are established in the guideline of mathematics, issued by the Ministry of Education, as follows:

Objectives

1. To identify the aim of the problem and to organize the data accordingly.
2. To classify, arrange and present data in tabular and graphical forms, and to compute means and ratios.
3. To read and interpret tendencies of data presented in tabular and graphical forms, and summarized in means and ratios.
4. To compare tendencies of the data with subsets of the data, and to explore the idea of inference.
5. To introduce the idea of probability as a measurement of the degree of uncertainty in stochastic events.

Plans in each grade

1st grade :

a. Identify the set having the same numbers of members as a given set.

b. Match a set with its number properly.

c. Classify figures by shape, colour, etc.

2nd grade :

a. Count the total number of figures, classifying them by shape, colour, etc.

b. Read a one dimensional table. Record data in a simple one dimensional table.

c. Read picture graphs showing comparisons and recognize their usefulness.

d. Construct simple picture graphs.

3rd grade :

a. Classify data by date, location etc. Display it in a table. Check missing and overlapping data by computing the total of the table.

b. Read and interpret a simple bar graph.

c. Display data presented in tabular form in a bar graph.

4th grade :

a. Classify and arrange data in a two dimensional table. Check missing and overlapping data by computing totals of the table.

b. Read and interpret a bar graph.

c. Round data to nearest tenth... the nearest hundredth, thousandth, and display it in a bar graph.

d. Read and interpret a broken-line graph (line graph). Display data in a broken-line graph.

5th grade :

a. Construct a table of proportions from a table of real numbers.

b. Display data in a band graph. Display data in a circle graph (pie chart). Read and interpret band and circle graphs.

c. Show what a mean means and use it correctly.

6th grade :

a. Compare histograms and introduce a measure of dispersion (range). Recognize that the mean and range summarise the tendencies of data.

b. Select and construct the most appropriate graph for a given data set among those studied in 6 years in elementary schools.

c. Explore the relationship of the mean and ratio of a sample with those of a population, and explore, informally, the idea of sampling with an emphasis on the representativeness of the samples.

d. Explore empirical probability by observing the outcomes of a simple experiment.

TABLE 1 : School Hours*

	Grade					
	1	2	3	4	5	6
Mathematics	136	175	175	175	175	175
Statistics**	3	4	7	12	6	17
% (math/stat)	2.2	2.3	4.0	6.9	3.4	9.7

* 1 school hour = 45 minutes, ** estimated

Table 1 lists school hours for mathematics and estimated school hours for statistics taught in mathematics. Table 2 summarizes the above plans. To characterize the plan we compared it to that of the U.S.A. Note that, in the U.S.A., each state is responsible for the education of its students; that there is no similar plan to regulate the teaching in all schools in the state as in Japan; and that a very flexible plan is generally established independently by each of the local school units, called school systems. In the table the above plan is compared with that established by the Chapel Hill-Carrboro School Systems (CCSS), North Carolina, which is a unit consisting of 6 elementary schools. Mathematics In Our World is the textbook which is used. The table shows that statistical topics taught in Japan are wide and extensive and that they are allocated to grades systematically, whereas the CCSS plan limits topics and teaches them repeatedly.

TEACHING OF STATISTICS IN MATHEMATICS TEXTBOOKS

A large portion of statistics taught in elementary schools in Japan consists of tables and graphs. They are formally introduced in the grades shown in the previous section. However, the prototypes of tables and graphs are abundantly and effectively utilized for introducing and developing mathematical ideas in earlier grades. This training, coupled with the recent development of colour printing technologies, facilitates students' quick and easy understanding of mathematics. Note that students are also exposed to tables and graphs in

TABLE 2 : Comparison of teaching plans of Japan and Chapel Hill-Carrboro School Systems (CCSS), North Carolina, USA

		Grade					
		1	2	3	4	5	6
1 dimensional table	Japan		0	0			0
	CCSS						
2 dimensional	Japan				0	0	0
	CCSS						
picture graph	Japan		0				
	CCSS	0	0	0	0		
bar graph	Japan			0	0		
	CCSS		0	0	0	0	0
broken-line graph	Japan				0		0
	CCSS			0		0	0
circle graph	Japan					0	0
	CCSS			0			
band graph	Japan					0	0
	CCSS						
histogram	Japan						0
	CCSS						
mean	Japan					0	
	CCSS					0	
dispersion	Japan						0
	CCSS						
sampling	Japan						0
	CCSS						
probability	Japan						0
	CCSS						

Note: CCSS textbooks deal with picture graphs in 5th and 6th grades, too. The case of single symbol representing a many to one ratio is discussed in upper grades. It also deals with mathematical probability of equally likely events in 6th grade. Textbooks in Japan deal with empirical probabilities which covers also unequally likely events.

---

science and social studies prior to being formally introduced in mathematics. Figure 1 displays a part of the results by Miyagawa (1981). She carried out a test in two elementary schools in Kusatsu City (the total number of students in each grade is about 200). The figure compares percentages of students correctly answering questions regarding tables and graphs in the one grade under and the grade where these topics were formally introduced. Although the correct answer rates depend on questions asked, the figure shows that students have a pretty good knowledge of tables and graphs before they are formally introduced in textbooks in mathematics. The finding gives rise to questions on the utility of the formal introduction (further discussion on the issue is given in the next section).

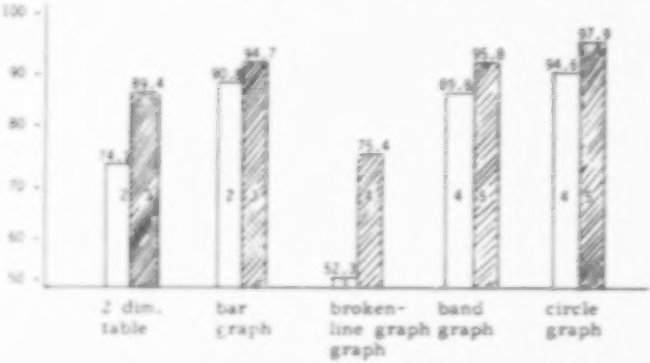
Figure 1

Percentages of students correctly answering questions regarding 2 dimensional table, bar graph, broken-line graph, band graph and circle graph.

▨ : the grade in which the table and graphs are formally introduced.

☐ : the grade immediately preceding the above.

Data source: Miyagawa, N. (1981)



A summary of examples showing how tables and graphs, or their prototypes, may be utilized effectively for introducing and developing important ideas in mathematics is shown below.

TEACHING OF STATISTICS IN ELEMENTARY SCHOOLS

We have seen in the previous section how effectively tables and graphs, and their prototypes, can be used for introducing and developing important ideas in mathematics. Similarly tables and graphs are effectively used for communicating information clearly and quickly in science and social studies. We, Japanese statisticians, encourage these teachings. For, so far as reading and interpretation are concerned, tables and graphs are easy and straightforward and, as was indicated in Figure 1, students today have pretty good ability to understand them. What is important and necessary for teaching of statistics is to give students experience in collecting data, summarizing it in tables, displaying

OBJECTIVES (MATHEMATICS)	TABLES, GRAPHS AND THEIR PROTOTYPES
	GRADE: 1
1. To identify the number of items in a group and to match the number names with the number they present.	To use picture graphs.
2. To compare two objects and be able to state which is larger or shorter; greater than, equal to, or less than etc.	... number lines. ... coloured tapes. ... picture graph.
3. To classify figures by means of shape or colour.	... 1 dimensional table.
	GRADE: 2
1. Understanding of addition, subtraction and how they are related.	To use number lines and coloured tapes.
2. Introduction of multiplication; showing how addition and multiplication are related.	... picture graphs.
3. To count the total number of figures by clasifying them by shape or colour.	... 1 dimensional tables.
	GRADE: 3
1. Introduction of the fractional parts (1/2, 1/3, 1/4).	To use circle and band graphs.
2. Introduction of addition and subtraction of fractional parts.	... number lines. ... coloured tapes.
3. Introduction of decimals: addition and subtraction of decimals.	... number lines. ... segments (coloured tapes).
	GRADE: 4
1. Introduction of round numbers: addition and subtraction of round numbers.	To use statistical data (populations of cities), tables and bar graphs.
2. Understanding the relationship of two dependent variables.	... broken-line graphs.
	GRADE: 5
1. Introduction and application of ratio and per cent.	To use band graphs and circle graphs.
	GRADE: 6
1. Introduction of direct proportion and inverse proportion.	To use broken-line graphs

it in graphs and interpreting its tendencies. The first objective of the guideline states "to identify the aim of the problem and to organize the data accordingly". Note that it says "organize" and not "collect". The collection of data is neither encouraged in the guideline nor in the textbooks, since it consumes school hours. However "to have a great reserve" is a recent slogan of education in Japan. We think that it is a good time to advocate incorporating problem solving in textbooks and allocating ample time for its study. In the problem solving we expect students to organize the various topics they have studied and to deepen their knowledge by solving problems taken from their daily lives. We should not overlook the importance of demonstrating to students, especially those in elementary schools, the usefulness of what they have studied at schools in their lives. Problem solving is frequently incorporated in the textbooks of science and social studies. We suggest that teachers should encourage students to apply statistical techniques to find facts in these subjects. Problem solving is very rare in mathematics. We suggest to establish it in mathematics, too. This is not only important for the teaching of statistics but also for the teaching of mathematics.

The other important topics in statistics which cannot be incorporated in the systematic teaching of mathematics are poorly handled in the guideline as well as in mathematics textbooks. In fact, the role of the mean for central tendency, dispersion, the idea of sampling and that of probability are introduced at the last part of the 6th grade, after almost all of the required elementary school mathematics have already been taught. We think it is unreasonable. The importance of these topics in our society has greatly increased. Students are exposed to these concepts in their daily lives and seem to have already the potential to understand them in earlier grades. In order to improve the teaching of these topics, we advocate the establishment of a section concerning mathematics in games in mathematics textbooks. Here, we may introduce the idea of random variation, the role of the mean for central tendency, the idea of dispersion and the idea of probability by discussing students' favourite games. In particular, our idea is to introduce them in the 4th grade, develop them in the 5th grade and then introduce the idea of sampling in the 6th grade.

The establishment of the topics of problem solving and mathematics in games must be, of course, well coordinated with all subjects taught in elementary schools. Further, consideration is necessary to reduce the chances of excessively increasing the students workload by incorporating these programs. For example, regarding statistics, we could reduce the workload of students a little by minimizing the formal introduction of graphs (for example, the formal introduction of band graphs could be eliminated) which are unnecessarily precise and uniform from the point of view of teaching statistics.

While the above suggestions are moderate in theory, it is still considered too radical to carry them out in practice in the present educational climate in Japan. Therefore, as the "next best" policy, we are encouraging teachers to give students experience in collecting and analyzing data in the following way:

(i) The school hours in Table 3 are reserved for special activities in elementary schools in Japan, and they can be used partly for solving problems arising from classroom activities/school lives of students. We are suggesting that teachers encourage students to look into these problems by applying statistical techniques they have studied.

(ii) As homework is regularly given to students during long (summer/winter) school holidays in elementary schools in Japan, we are encouraging teachers to give students problems in collecting data such as recording daily temperature etc., displaying the data in tables and graphs, and interpreting the data's tendencies.

The suggestions depend totally upon the teachers' enthusiasm and their knowledge of statistics. Unlike mathematics, the answer to a question in statistics is not necessarily unique and fixed. This requires that teachers have sufficient abilities in teaching statistics. However, the problem is that not many teachers in Japanese elementary schools have such abilities. As teachers have several workshops on teaching each year, we have strongly encouraged them to establish workshops on teaching statistics. So far as teaching statistics in high schools is concerned, quite a few workshops have been successfully established and progress has been made with the consultation of statistical scholars. However, almost no statistical teaching workshops have been established so far in elementary schools. We find that obtaining the cooperation of teachers in elementary schools is a very difficult task. We hope that the recent developments in our environment shown in the introduction can change it.

TABLE 3 : School Hours for Special Activity

Grade	1	2	3	4	5	6
School hours*	34	35	35	70	70	70

* 1 school hour = 45 minutes

Acknowledgements

The author is indebted to Professor Shiro Yamazoe for his critical discussions during the preparation of this monograph. He also expresses his deepest appreciation of Mr. Carl Yoshizawa for his generous help in refining the English of this manuscript.

References

[1] Chapel Hill-Carrboro School System (1981) Mathematics Objectives.

[2] Ministry of Education (1977) Summary of Guidelines for Elementary Schools, (in Japanese).

[3] Ministry of Education (1978) The Guideline for Elementary Schools; Mathematics, (in Japanese).

[4] Miyagawa, N. (1981) Teaching of tables and graphs in elementary schools and its consideration, B.A. dissertation, School of Education, Shiga University (in Japanese).

[5] Mathematics In Our World, (1981) Addison Wesley.

TWO EXAMPLES OF STATISTICS ACTIVITIES FOR ELEMENTARY SCHOOL CLASSES

Clare Heidema, CEMREL-CSMP, St.Louis

SHUNDA'S NEWSSTAND

The business world offers many opportunities for statisticians, especially in advising decision-making processes. Elementary school age youngsters in their everyday lives also are confronted with decision situations, some with a business flavor, where a rudimentary understanding of statistics may prove useful. However, the pedagogical concerns of making the study of these situations accessible at an early age often prohibits consideration until the problems can be discussed on a high mathematical level. Consider, for example, the classic "Newsboy's Problem" concerning a newspaper seller attempting to maximize profit. Children can appreciate such a problem; indeed they may have paper routes or operate newspaper stands themselves. SHUNDA'S NEWSSTAND provides an excellent example of how an operations research problem involving statistics in a decision-making context can be presented at the elementary school level.

Crucial to the presentation of the problem in terms that 9-11 year olds can understand is the use of stimulating pictures to view the sample data in a variety of ways. The pictures and graphs provide an alternative to technical numerical methods; an alternative that is both aesthetically and pedagogically appealing.

Shunda's Problem

A common problem in the business world is to determine the optimal quantity of items to buy or produce in order to make the most profit. Many factors affect such decisions and no one can be guaranteed of always making the best choice. Still, educated decisions can be made by doing a careful study of sample data and by assuming that past behavior of the consuming public is a good predictor of future behavior. There is always a risk; all one can do is decide what is the most reasonable prediction. SHUNDA'S NEWSSTAND is concerned with such a business world problem as it affects Shunda, a young newspaper seller.

We present the situation in a story-workbook*, that is, as a story told in the pages of a "comic book". Along the way the students respond to questions and solve intermediate problems. They become familiar with a variety of pictorial representations of information and supply the necessary results needed to understand the main problem. Prompted by a series of stimulating pictures for recording data, the students use the daily demand to determine the most profitable inventory of newspapers.

---

* SHUNDA'S NEWSSTAND is in the CSMP Library, a collection of math story-workbooks providing fanciful excursions into the colorful world of mathematics.

Our story is about Shunda, a newspaper seller, who has a newsstand on the corner of Hamilton Street and Euler Avenue. Each day between 4:00 and 6:00 PM, she sells newspapers to people passing by her stand.

Several pages at the beginning of the booklet* present basic information about Shunda's job.

Shunda buys newspapers from a dealer for 10 cents each, and she sells them to her customers for 20 cents each. Since the number of Shunda's customers varies from day to day, the dealer agrees to buy back the unsold papers for 5 cents each.

Shunda wants to have a successful business and to make as much money as possible. She devises procedures for keeping records of her business from day to day. The students work through several pages of practice using Shunda's record-keeping procedures. Selected pages from the booklet will suggest the variety of pictures and the nature of the individual work.

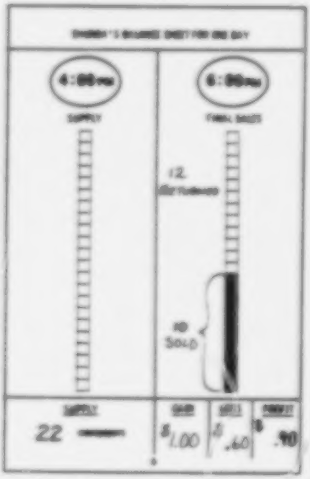
---

* In this paper, we use actual reproductions (in reduced form) from the booklet, omitting color. Light and dark shading will replace essential distinction of red and blue.

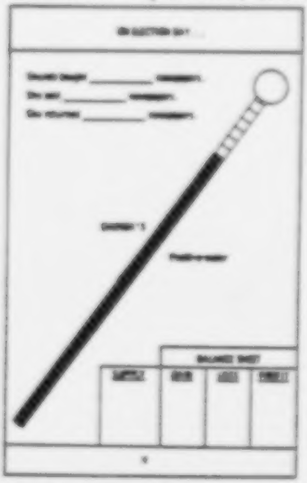
BEST COPY AVAILABLE

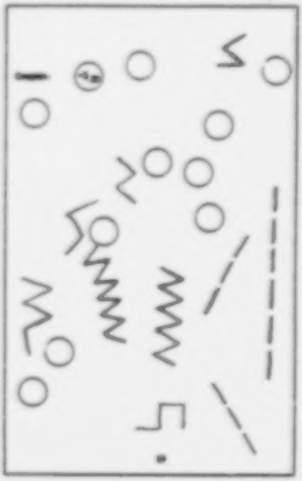
BEST COPY AVAILABLE

Shunda begins her newspaper business as an apprentice, so may buy differing numbers of papers from day to day depending on what she expects the demand to be.

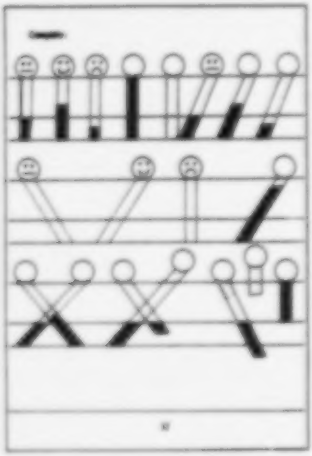


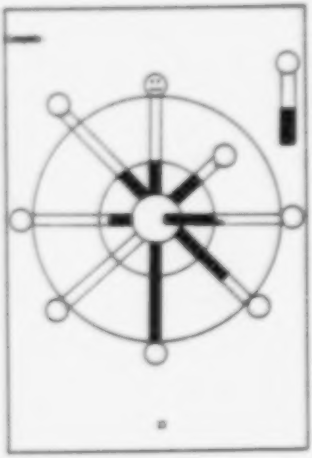
Stylization on page 10 reflects Shunda's whimsical side. Such artistic freedom requires little or no explanation, and it contributes to the students' experience viewing data in a variety of ways.





On days when Shunda has a positive profit, she is happy; when she has a negative profit, she is unhappy; and when she breaks even, she shrugs her shoulders and hopes for a better next day.

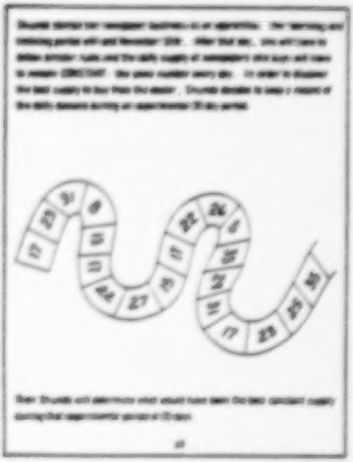


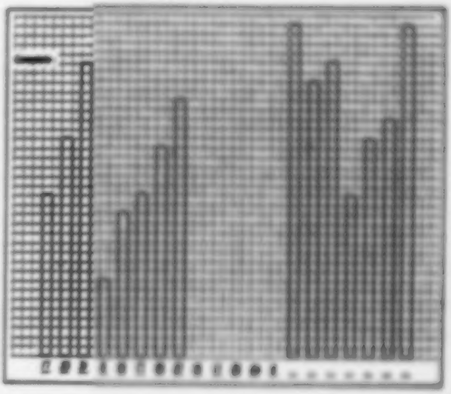


Students must use only the light:dark ratio to determine Shunda's mood or sales on several days. This step suggests the critical role in Shunda's business of a one-third, two-thirds ratio of returns to sales. Later we will see how these pictures contribute to a way of determining a most profitable inventory.

The story continues with new information that presents Shunda's main problem. The training period for Shunda will end shortly. Then she will have to follow stricter rules; the daily supply of newspapers will have to remain constant; the same number every day. Shunda must decide what constant daily supply would be the best to buy from the dealer. How can she make such a decision? In order to determine the best daily supply of newspapers to buy from the dealer, Shunda keeps a record of the daily demand during an experimental twenty-day period.

Students draw a graph of Shunda's sales record:





Shunda hopes that if she can determine what would be the best constant supply during that experimental period, then she can expect it also to be the best constant supply in the future.

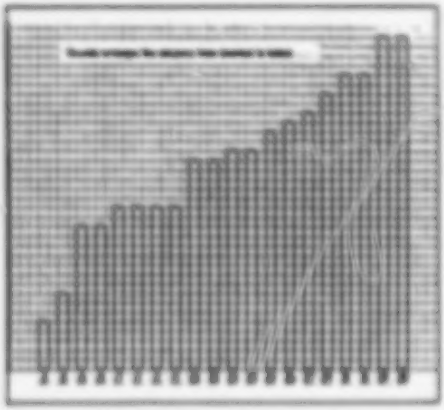
Before proceeding towards a solution, students stop to express their opinions on Shunda's decision. Such a problem interests 10 to 12 year olds; they may be just beginning to earn their own spending money. This discussion can bring out the students' natural curiosity about business matters; it tests their intuition about implications in the statistics for a twenty-day experimental period. Our experience suggests that students will think of a variety of ways of viewing the data; for example,

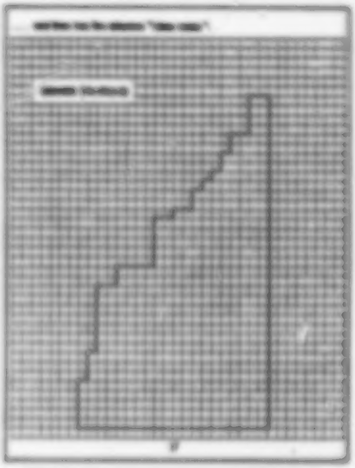
- Shunda should get only as many papers as she is sure to sell; then she is sure to have a positive profit every day.
- Shunda should get about 17 newspapers every day because she most often sold that number of papers.
- Shunda should get about 22 newspapers every day because half the time she sold more than 22 newspapers and half the time she sold less than 22 newspapers.
- Shunda should get about 20 newspapers because that is halfway between 35 (the most she sold) and 5 (the least she sold).
- Shunda should get 22 newspapers every day because that is the average number she sold daily during the experimental period.

BEST COPY AVAILABLE

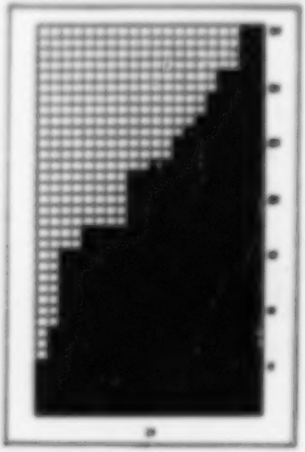
Shunda's Solution

We continue our story with Shunda organizing and studying the demand process as reflected in her records for the experimental twenty-day period.





The largest demand during the twenty-day experimental period was 35. Suppose Shunda would decide on a constant daily supply of 35 newspapers. Page 19 of the booklet shows a record for the twenty-day experimental period with the constant supply of 35 newspapers. We can view this picture as putting Shunda's profit-o-meters for the twenty days side by side and as highlighting the demand staircase.



Students complete a balance sheet . . .

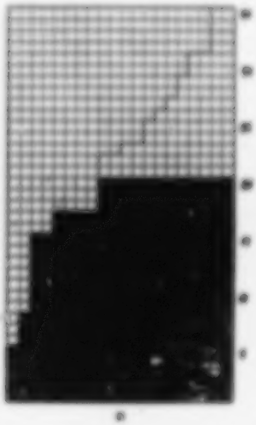
BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
35	$44.00	$15.00	$31.00

BEST COPY AVAILABLE

and then question the choice of 35 as a constant daily supply. With such a supply, Shunda would satisfy her customers every day, but is it her most profitable choice? It seems that there is considerable loss (light) in the previous picture.

With the above picture and previous experience of finding daily gain, loss, and profit, students explore "what if" questions in preparation for determining a most profitable constant daily supply.

Suppose Shunda would decide on a constant daily supply of 20 newspapers. Page 21 of the booklet shows a record for the twenty-day experimental period with the constant daily supply of 20 newspapers. Would this choice produce more or less profit than a constant daily supply of 35 newspapers? There is much less loss (light) but also less gain (dark).

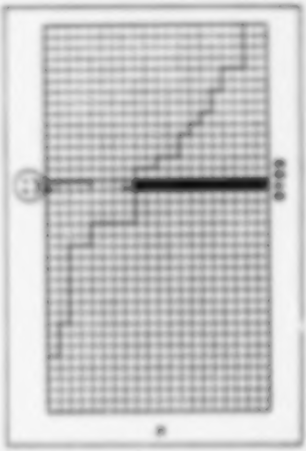


Students complete a balance sheet . . .

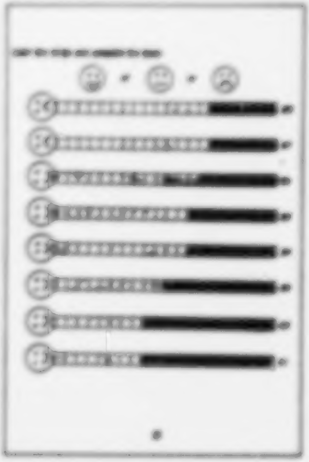
BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
20	$35.10	$2.45	$32.65

and find that 20 is a better choice than 35 for the constant daily supply. But is it the best possible choice?

The next two pages of the booklet suggest a way to view the effect of changing the constant daily supply. Therefore, with a constant daily supply of 22, Shunda's total profit would be $34.25 ($32.65 + $1.60). This procedure demonstrates how the earlier stylizing of Shunda's profit-o-meter contributes appropriate experience for studying the graph that describes the demand process.



Students complete additional profit-o-meters on pages 24 and 25 of the booklet (shown here with solutions) to find that Shunda could continue to increase her profit by increasing the constant daily supply until the supply is 26. Changing the constant daily supply from 26 to 27 would result in less profit as suggested by the frown on the horizontal line for 27.



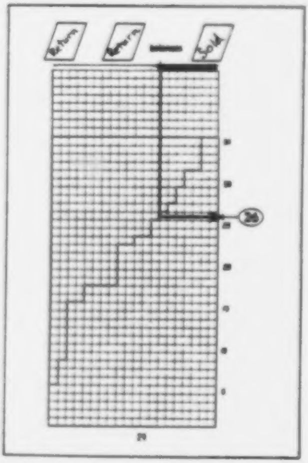
BEST COPY AVAILABLE

Shunda's Conclusion

During the twenty-day experimental period, the largest profit would have been obtained with a daily supply of 26 newspapers.

BALANCE SHEET for 20 day experimental period			
DAILY SUPPLY	TOTAL GAIN	TOTAL LOSS	TOTAL PROFIT
26	$40.80	$5.60	$35.20

Returning to the early observation that one newspaper sold balances two newspapers returned, Shunda takes another look at the graph describing the demand process for the twenty-day experimental period.



BEST COPY AVAILABLE

The story-workbook ends with a discussion of Shunda's conclusion. Since 26 papers would have been the best constant supply for the twenty-day experimental period, Shunda decides to adopt it. Of course, she cannot be sure this will remain the most profitable choice. She does not try to forecast future demand; what she does is make the best choice based on the limited evidence of the experimental period. As in most business ventures, she cannot avoid taking some risk.

This example of statistics activities for the elementary school intends to involve students in a real world application of mathematics. There is considerable calculation practice in a meaningful context. There is the opportunity to use bar graphs in a dynamic way to organize data. And, most importantly from a mathematical point of view, there is the experience of using statistics for problem solving.

POPULATION GROWTH

Population statistics appeal to students' natural curiosity about the world around them and the future that lies ahead. The sixth-grade CSMP curriculum includes a series of lessons on population growth. The lessons provide the opportunity to organize, interpolate and analyze real life statistics. More importantly, students use the data to make inferences about the past, present, and future - a rare activity at the elementary level in the study of statistics.

Population Growth Rates

A first lesson introduces students to the concept of population growth rate. Using their intuitive sense of ratio, students develop techniques to compute the growth rates and net population gains of several United States' cities.

The lesson begins with a discussion focused on the question: "What factors affect population growth?" Students conclude that the factors affecting population fall into four categories: births, deaths, immigration, and emigration. They proceed to develop several methods for computing growth rate and net gains. By not providing specific procedures for computing the net growth or the growth rate, the teacher allows students to create their own techniques. The strategies they develop increase their understanding of the concept "population growth rate" - the key for later lessons in this sequence.

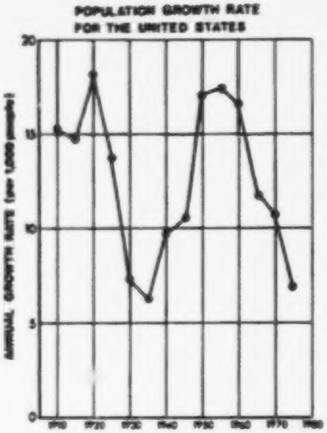
Tables and Graphs

The activities in a second lesson demonstrate that appropriate organization of data facilitates both the interpretation of the data and inferences based upon the data. Students begin by examining a mock newspaper article:

> The United States population growth rate has been declining. Growth rate per 1,000 people was 17.5 in 1955 whereas it was 10.7 in 1970 and 7.0 in 1975. It also declined for a while before World War II; it was 13.8 in 1925 and only 6.3 in 1935. There was a sharp increase in the growth rate after this war increasing from 10.7 in 1945 to 17.1 in 1950. Other yearly growth rates were 15.2 in 1910; 14.9 in 1915; 18.8 in 1920; 7.3 in 1930; 9.8 in 1940; 16.6 in 1960; and 11.8 in 1965.

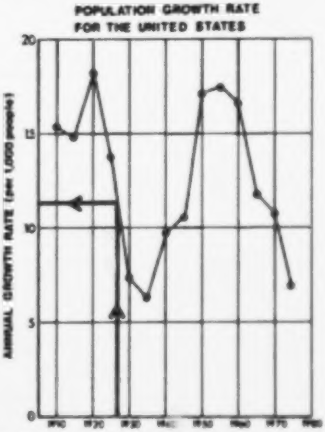
The teacher highlights the effect of the disorganized nature of the data by asking several questions such as: When was the growth rate the highest? lowest? During what periods was it increasing? declining? The students' difficulty in answering these questions motivates the central theme of the lesson: Are there better ways in which to present this data? After constructing a table and a graph, the class reconsiders the questions they previously found difficult thereby experiencing the benefit of organized data.

Year	Annual Growth Rate (per 1,000 people)
1910	15.2
1915	14.9
1920	18.8
1925	13.8
1930	7.3
1935	6.3
1940	9.8
1945	10.7
1950	17.1
1955	17.5
1960	16.6
1965	11.8
1970	10.7
1975	7.0



Further questions reveal the advantage of the graph over the table. For example, what was the growth rate in the year 1927? Through interpolation, students can estimte a 1927 growth rate of 12 people per 1,000. The graph also shows the large fluctuations in the U.S. population growth rate.

BEST COPY AVAILABLE



The remainder of the lesson focuses on some historical events that could have caused fluctuations. Students suggest the Great Depression and World War II as possible causes for the low rate of growth from 1930 to 1945. However, changing sociological views, such as people deciding to marry later or electing to have fewer children, seem a more likely cause of the recent low growth rate from 1970 to 1975.

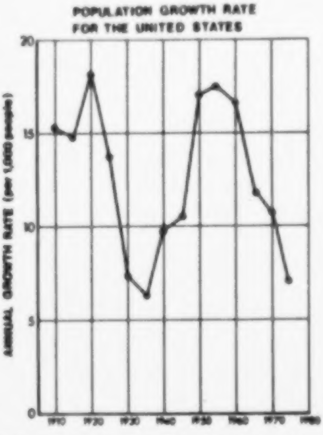
The activity of organizing data, found in the first part of this lesson, is fairly commonplace in the study of statistics. However, the discussion of factors that may have caused the large fluctuations in the population growth rate represents an important shift in the lesson. The students move from simply reading the data to interpreting and making inferences based upon the data. These attempts to interpret data allow students to appreciate the purpose of organizing data into graphs, rather than rotely practising techniques of data organization. Correlating data with known historical and sociological factors provides a framework for later using current data to predict future events.

Population Pyramid

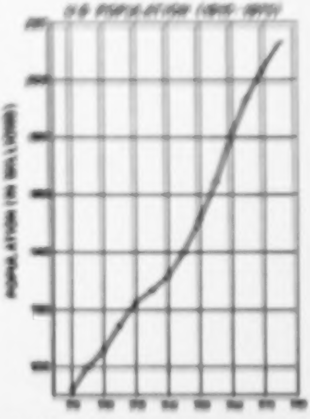
There are many ways to graph a set of population data. The choice of a particular graph reflects the feature of the data that the statistician wishes to highlight. In a third lesson, students explore several population graphs. First, the comparison of two distinct graphical representations of population growth during the period 1910 to 1975 allows students to observe the different ways that each graph portrays the same information.

A: Population Growth Rate for the U.S. (1910-1975)

B: U.S. Population (1910-1975)



BEST COPY AVAILABLE



Since these two graphs display the same census data, they reflect the same trends in different ways. For example, the population is always increasing, shown in graph A by the rate always being positive and in graph B by the total population always rising. Also, the low period in growth from 1930 to 1940 shown in graph A corresponds to a "leveling off" in graph B. Conversely, the sharp rise from 1950 to 1960 in graph B indicates the high growth rate that is recorded in graph A. This activity encourages students to observe different ways that graphs can display rates of growth. In particular, the graphic comparison demonstrates that a declining (but positive) growth rate and a rising total population can occur simultaneously.

Besides studying the overall growth rate, demographers also analyze the United States population according to various factors: age, race, sex, religion, etc. With this in mind, the following statistic motivates the idea of population distributions.

United States: K-8
1970: 34,300,000 students
1975: 32,000,000 students

The decline in total elementary school enrollment surprises most students. Who might be concerned about a decline in school enrollment? Teachers are most certainly concerned. Implications of this single statistic include school closings, teachers being laid off, and many other all too familiar problems. An open student discussion of this statistic focuses on such implications.

When first confronted with the above statistic, students might assume that the decline in school enrollment is due to an overall decline in population, contradicting the previous observation that the United States had a positive growth rate from 1950 to 1975. We then present the idea of population distribution by age group.

U.S. Population by Age Group – 1976

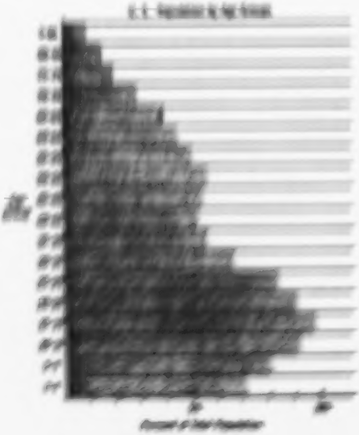
Age	Percent of Total Population
0-4	7.1
5-9	8.1
10-14	9.2
15-19	9.8
20-24	9.1
25-29	8.3
30-34	6.6
35-39	5.5
40-44	5.2

Age	Percent of Total Population
45-49	5.4
50-54	5.6
55-59	5.0
60-64	4.3
65-69	3.9
70-74	2.8
75-79	1.9
80-84	1.3
>84	0.9

By comparing the percent of population in the 5-to-14 age groups (the approximate K-8 enrolment in 1975) with the 10-to-19 age groups (the K-8 enrolment in 1970), students note that the enrolment loss appears due to a decrease in the number of people in that age group. The table reveals how the population in a particular age group can decline even as the total U.S. population continues to increase.

The introduction of population distribution by age suggests other questions of human interest such as: How many Americans remember World War II? World War I? How many Americans are over the age of 65?

After exploring these questions, we conclude the lesson by constructing a population pyramid of the United States population in 1976.



BEST COPY AVAILABLE

Comparative Population Graphs

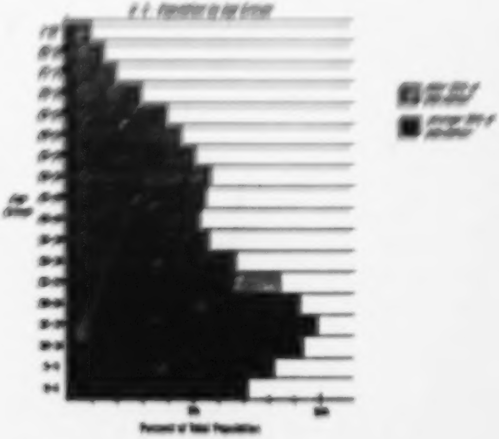
Upon graphing and interpreting a population pyramid for the United States, students naturally assume pyramids for other countries to be "normal". In a fourth lesson, the very dissimilar shapes of both Sweden's and Mexico's population distribution graphs as compared to the United States' graph disagrees with the students' assumption. A goal of this lesson is to interpret the three countries' graphs as a means for conjecturing the political priorities of the countries and what problems may lie in the future for each nation.

The beginning of the lesson returns to a conjecture made in the first lesson. Between the years 1930 and 1945 the United States experienced a low rate of growth. Two historical occurrences were given as possible causes: The Great Depression and World War II. The teacher challenges the students to use the population pyramid to determine which event appears to have had greater impact.

A notable feature of the US population pyramid is the relatively low percent of people in the 40 to 44 age group compared to neighboring age groups. Being of age somewhere between 40 and 44 in 1976 means being born sometime in the years from 1932 to 1936 - the midst of the Depression. Apparently, either people during the Depression chose to have fewer children or their children had a much lower life expectancy. The graph suggests that the Depression was a greater contributing factor to the lower population growth rate than was World War II.

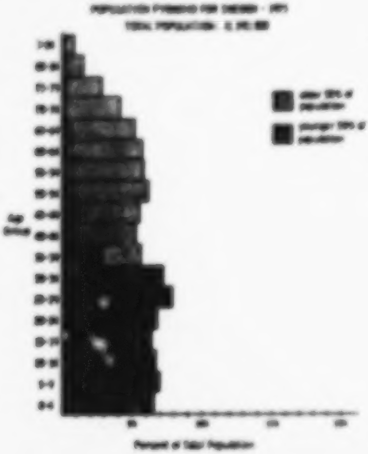
Besides graphs, another simple statistical tool for comparing national populations is averaging. The population pyramid does not allow easy computation of the mean age of the population; however, the median age is appropriate and is easily computed.

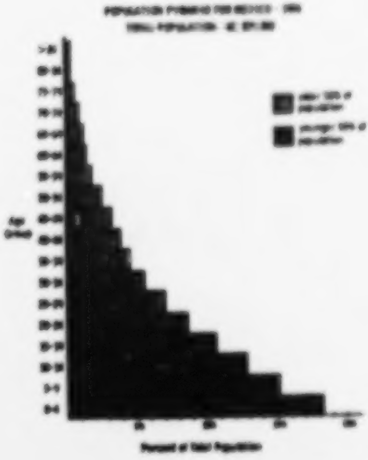
The median age of the US population in 1976 was about 28.



BEST COPY AVAILABLE

After completing this computation, students discuss whether other countries are likely to have similar median ages. Even at first glance, the population pyramids for Mexico and Sweden reveal dramatically different population distributions in different countries.





BEST COPY AVAILABLE

With Sweden's uniform population distribution, its median age is about 35, higher than that of the United States. In Mexico, a greater portion of the population is in the younger age groups; the country has a low median age of about 17.

What insights into areas of concern for these governments do these population graphs and medians suggest? The remainder of the lesson focuses on a comparison of some political and social issues these countries may soon face, such as:

- United States and Sweden have a large population in the older age groups who need support.
- United States is closing schools while Mexico needs to build more schools.
- Mexico's rapidly increasing population could cause shortages of food, housing, and health facilities.
- Sweden has the largest percent of older people, while Mexico has the smallest.

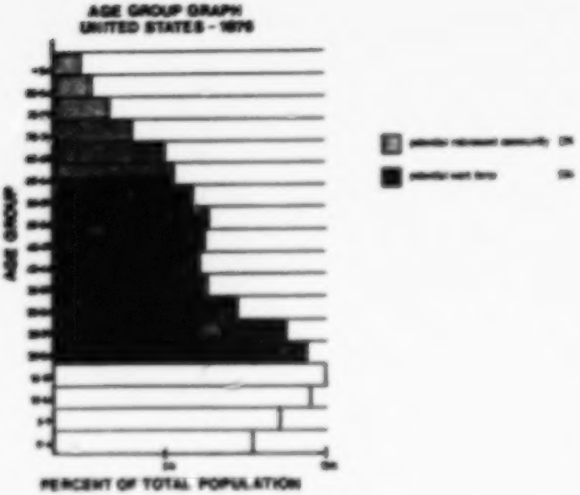
As in previous lessons, the lesson has moved to a discussion about the data. The importance is not so much the statistical tools that the students acquire, but the important realization that these tools enhance the discovery and discussion of events in the world around them.

An extension of the lesson studies the impact of the current population distribution of the United States to a topic of recent controversy: the Social Security System.

There have been several rises in the Social Security tax in recent years, each one promising to set the system on firm ground. As we shall see, the problem with the system may lie in its basic design, therefore calling for more creative measures to rectify the system. The system is designed so that the current work force generates the revenue for the retirees currently on social security. In order to gain only a crude overview and to simplify the analysis, we will contrast the total potential work force (ages 20-62) to the total potential social security receivers (over 62).

The following questions highlight some difficulties in the Social Security System:

- How many people in the potential work force does it take to maintain one person on social security?
- In 1981, how many people entered the work force for every person who entered the retirement system?
- Will the situation get better or worse? in 1980? (Assume life expectancy of 80).



The graph provides answers to these questions. In 1976, approximately four people (4 x 13 is about 53) in the potential work force were needed for every one person on social security (a ratio of 4 to 1). In 1981, however, approximately 5% of the population entered the social security system while only 10% of the population entered the work force (a ratio of 2 to 1); thus the potential work force is not growing as fast as the potential retirees. And as we can see from the graph, the problem will continue to worsen in the future since the population entering the work force is decreasing. The important statistical activity again gives the students an opportunity of not just taking the data at face value but drawing implications from the data in regard to some social issues.

BEST COPY AVAILABLE

SUMMARY

The series of lessons on population growth exhibits several valuable features for the teaching of statistics, namely:

- topics that interest students at a particular grade level,
- a problem-solving atmosphere, and
- a unifying factor for a variety of mathematical tools and concepts.

Students' curiosity about the world around them begins to emerge in the intermediate grades. Utilizing this curiosity is valuable in teaching, in particular teaching mathematics, effectively. Building on a foundation of interest, the mathematical content is then introduced as a means to resolve questions and to generate a source of new inquiries, thus creating a true problem-solving atmosphere. Students learn new statistical methods as well as gain an immediate appreciation of their application.

References

I. The story of SHUNDA'S NEWSSTAND was written by Georges Papy. Clare Heidema developed the series of lessons in the CSMP curriculum based on this story-workbook and she also wrote the paper on this topic in the book, CSMP: Probability and Statistics.

II. Richard Armstrong developed the series of lessons on population statistics for the CSMP curriculum and Thomas Giambrone wrote the paper on this topic in the book, CSMP: Probability and Statistics.

USING GAMES TO TEACH SOME PROBABILITY CONCEPTS

*George W Bright, Northern Illinois University, John G Harvey, University of Wisconsin
and Margariete M Wheeler, Northern Illinois University

Since 1976 the authors have been engaged in a systematic investigation of the cognitive effects of games on mathematics learning. This investigation was initiated and continues for a variety of reasons.

1. Games seem to be important instructional activities, in the sense that teachers can choose to use a game in a particular instructional setting for instruction pertaining to a particular instructional objective.

2. Teachers of students aged 6-11 report (Graeber, Rim & Unks, 1977) regular use of games in their classrooms. The types of games used were not surveyed.

3. Students aged 6-11 report (National Assessment of Educational Progress, 1979) not only that mathematics instructional games are used in their classrooms but also that games seem to help them to learn mathematics.

4. Prior to 1976 the limited and very fragmented research on the cognitive effects of mathematics instructional games had produced few substantiated results (Bright, Harvey & Wheeler, 1977).

From the outset, this investigation has focused on identifying effective uses of games so that recommendations for appropriate incorporation of games into school practices can be made. A game is defined by seven criteria:

1. A game is freely engaged in.
2. A game is a challenge against a task or an opponent.
3. A game is governed by a definite set of rules. The rules describe all procedures for playing the game, including the goal(s) sought; in particular, the rules are structured so that once a player's turn comes to an end, the play just completed cannot be retracted or exchanged for another move.
4. Psychologically, a game is an arbitrary situation clearly delimited in time and space from real-life activity.
5. Socially, the events of the game situation are considered in and of themselves to be of minimal importance.
6. A game has a finite state-space (Nilsson, 1971). The exact states reached during play of the game are not known prior to beginning of play.
7. A game ends after a finite number of moves within the state-space.

The first five of these criteria, based upon an earlier discussion of play, games, and gaming situations (Inbar & Stoll, 1971), give formal, psychological, and sociological characteristics of games. The last two criteria were evolved in order to exclude other, related activities or structures such as play and puzzles.

An instructional game is a game for which a set of instructional objectives has been determined; these instructional objectives may be cognitive or affective ones and are determined by the person(s) planning instruction. The source of ideas for the instructional games used in the research was Developing Mathematical Processes (DMP) (Romberg, Harvey, Moser & Montgomery, 1974, 1975, 1976). DMP is a rich source of mathematics instructional games for elementary school students. These games frequently also seem easily adoptable for other ages and for more sophisticated content.

Before the details of the research are presented, the concept of cognitive effects needs to be defined. In describing a taxonomy of educational objectives, Bloom (1956) stated that the cognitive domain "includes those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (p.7). Thus, cognitive effects are effects which describe changes or differences in the attainment of objectives in the cognitive domain. For the study reported here, the cognitive effects are defined as the change in attainment of an objective in the cognitive domain due to a particular instructional treatment; that is, due to the playing of mathematics instructional games.

In DMP one of the grade six (about age 11) lessons is designed to introduce the mathematical concept of fair/unfair by having students play three pairs of related games. (One pair of the games follows the reference list). For each pair, play of the games results in counts of the number of times each of the two players is a winner. By comparing these counts, students decide which game of the pair is more fair. The three pairs of games were expanded to eight pairs (Bright, Harvey & Wheeler, 1981) and used in a series of studies, one of which was with 36 grade six (about age 11) students.

The assumptions underlying this study included the following:

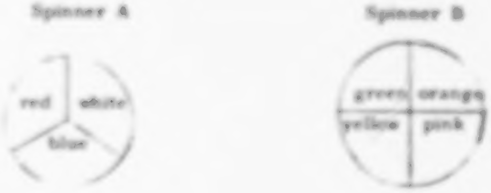
1. The mathematics content was substantial and would not be learned without instruction specifically designed to teach that content.
2. Prior to the start of the study, the students had not received instruction intended to provide mastery of the content.
3. To measure the effectiveness of games on substantial content, which were played in isolation from other instruction on that content, it was not important to have a control group which did not play the game.
4. The concept, fair/unfair, as embodied in the games was at the knowledge cognitive level (Bloom, 1956). Knowledge cognitive level content includes knowledge of specifics or of ways and means of dealing with specifics (pp.65,68).

The study was conducted on ten non-consecutive school days, during five consecutive weeks, in schools in Illinois. On day 1 two pretests were given: the four-card test of formal operations (Adi, Karplus & Lawson, 1980) and a 28-item content test. In each of the 28 items two games were described and students were asked to determine the fairer of the pair. For one 14-item subscale, items were accompanied by data summarizing 50 plays of each of the two games. (These data were generated by computer simulations of the games). The items on the other 14-item subscale were not accompanied by simulated plays. A sample item from the second subscale is provided in Figure 1.

Figure 1

Sample Test Item for Fair/Unfair Games

Which game gives you the best chance of winning or doesn't it make any difference?



BEST COPY

A. Game 1 You spin Spinner A once and you win if you get blue.

B. Game 2 You spin Spinner B once and you win if you get green or orange.

C. It doesn't make any difference.

On days 2 through 9 the eight pairs of fair/unfair games were played. A different pair was played for 20 minutes on each of these instructional days.

After playing each pair, students were asked to choose the fairer game, or to say that the games seemed equally fair. On day 10 the content test was given as a posttest.

The means and standard deviations and the comparisons of the pretest and posttest scores are presented in Table 1. In order to determine whether there were sex-related differences, the posttest scores for each subscale and for the total test score were used in an analysis of covariance with the four-card test and the corresponding pretest score as the covariates and with sex as the classification variable. None of these analyses revealed significant sex-related differences.

Table 1

Pretest - Posttest Comparisons

Instrument	Pretest Mean (SD)	Posttest Mean (SD)	t	p
14 items with data given	5.72 (1.70)	7.33 (1.74)	4.33	.000
14 items without data given	5.31 (1.67)	5.50 (1.88)	0.45	.654
total test	11.03 (2.56)	12.83 (2.66)	3.14	.003

The fair/unfair games seemed to have helped students learn to distinguish fair from unfair situations in the presence of data about the outcomes of playing those games. This effect did not seem to generalize to situations in which students were presented two games without any accompanying data. Perhaps this result is related to the fact that students had not received instruction on probability or on the concept of fair/unfair. Their only source of such knowledge would have been their personal experiences outside the classroom.

It is encouraging that the students seemed to have learned from their interactions with each other about how to interpret the data that they generated by playing the games. This same pattern of results was also noted in a replication of this study with grade eight students (Bright, Harvey & Wheeler, 1982).

One possible follow-up study would be to interview students regarding determination of the fairer game when sample data are not provided. The effects of knowledge about probability might be expected to display itself in the strategies that students use in making this determination. For example, examining students' perceptions about relative areas of spinners or relative distributions of sums of numbers on dice might provide useful information about the ways that students try to build mathematical structures from their personal knowledge bases.

References

[1] Adi, H., Karplus, R. & Lawson, A.E. (1980) Conditional logic abilities on the four-card problem: Assessment of behavioral and reasoning performances. Educational Studies in Mathematics, 11, 479-496.

[2] Bloom, B.S. (Ed.) (1956) Taxonomy of educational objectives. New York: McKay.

[3] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1977) *Cognitive effects of games on mathematics learning*. Madison, WI. University of Wisconsin - Madison. (ERIC Document Reproduction Service No. ED 166 007).

[4] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1981) Fair games, unfair games. In Schulte, A. & Smart, J. (Eds.), *Teaching statistics and probability: 1981 yearbook of National Council of Teachers of Mathematics*. Reston, VA: NCTM, 49-59.

[5] Bright, G.W., Harvey, J.G. & Wheeler, M.M. (1982) *Teaching statistics concepts through instructional games*. Paper presented at the First International Conference on the Teaching of Statistics, Sheffield, England, August 1982.

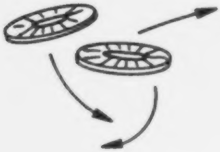
[6] Graeber, A.O., Rim, E. & Unks, N.J. (1977) *A survey of classroom practices in mathematics: Reports of first, third, fifth and seventh grade teachers in Delaware, New Jersey and Pennsylvania*. Philadelphia, PA: Research for Better Schools, Inc.

[7] Inbar, M. & Stoll, C.S. (1970) Games and Learning. *Interchange*, 1(2), 53-61.

[8] National Assessment of Educational Progress (1979) *The second assessment of mathematics, 1977-78, released exercise set*. Denver, CO: The author, May 1979.

[9] Nilsson, N. (1971) *Problem solving methods in artificial intelligence*. New York: McGraw-Hill.

[10] Romberg, T.A., Harvey, J.G., Moser, J.M. & Montgomery, M.E. (1974, 1975, 1976). *Developing mathematical processes*. Chicago: Rand McNally.

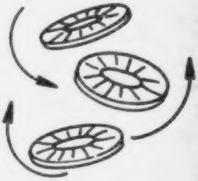


Flipped Chip

Game 1

(2 players)

Each player will need 15 chips. Take one extra chip and put a blank label on each side. Label the sides A and B. Decide which player will be A and which will be B.



Flip the chip. If A comes up, player A places a chip on his or her car. If B comes up, player B places a chip on his or her car.

The winner is the first person to fill all 15 chip spaces. Play two or three games. Then answer these questions:

Does each player have an equal chance at winning?
Does the same person win each time?
Is the game a fair game?

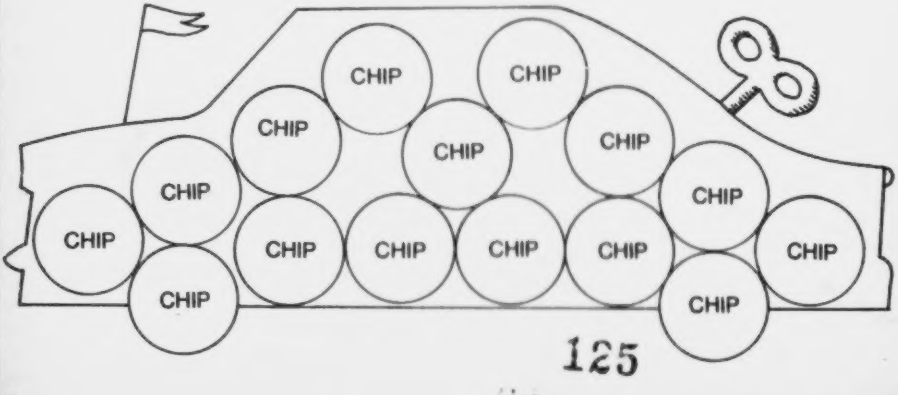
Game 2

(4 players)

Label three chips so that each has an A side and a B side. One player (AAA) puts a chip on his or her car if three A's come up. Another player (ABB) puts a chip on his or her car if exactly two B's come up. The third player (AAB) puts a chip on his or her car if exactly two A's come up. The fourth player (BBB) puts a chip on his or her car if three B's come up.

Flip all three chips at the same time. The winner is the first player to cover his or her car. Play two or three games. Then answer these questions:

Does each player have an equal chance at winning?
Does the same person win each time?
Is the game a fair game?



TEACHING STATISTICS IN ELEMENTARY SCHOOLS IN JAPAN

H Midzuno, Nippon Institute of Technology and Prime Minister's Office, Japan

Before starting a discussion of Statistics Education (SE) at elementary schools, it may be necessary to explain the nature of elementary education in my country.

Nine years of elementary education is compulsory. Parents or custodians of children between 6-14 years of age are obliged to send their children to schools, as stated in the Basic Law of Education. The Law is considered as a supplement to the Constitution, with a special preamble referring to the Constitution.

In the Constitution the basic right to receive education and the basic obligation of parents and custodians to provide education are stated with some related matters but the details are delegated to the Law.

Article 26 of the Constitution states:
"All citizens have the right to receive education equally, according to their ability as specified in the Law. All citizens are obliged to let the children under their custody receive ordinary education as specified in the Law. Compulsory education shall be free".

The preamble to the Basic Law of Education states:
"We have established the Constitution of Japan and revealed our determination that we construct a democratic and cultural nation and contribute to world peace and the welfare of the human kind. The realization of this ideal has to base itself upon the power of education.
We expect to bring up persons who respect the dignity of the individual and seek for truth and peace, as also we have to promote thoroughly education aiming to create a culture which is both universal and unique.
Expressing clearly the objective of education based on the spirit of the Constitution of Japan, we hereby enact this Law, in order to establish the basis of education for the new Japan."

Based on this, the Law determines many points important in education; e.g. objective and so forth, and as the result of these provisions the basic nature and scheme of education have been determined. Article 6 states "Schools referred to in the Law are public in nature and shall be established either by the nation, local public bodies or the legal bodies stipulated in the Law". Objectives are stated and matters pertaining to educational courses are to be determined by the supervisory organs.

An important specification is about the textbooks. Only textbooks approved by the Minister of Education or prepared by the Ministry of Education can be utilized.

Accordingly the substance is included in the Guideline. This is issued by the Ministry of Education, as the supervisory organ is set to be, for the time being, the Ministry of Education.

Thus the Guideline prepared and the textbooks approved determine the contents of teaching almost entirely.

The Guideline was revised in 1978, and education has been going on since 1980 using thoroughly the new materials.

The Composition of the educational courses is set to be,

National language
Social discipline
Mathematics
Science
Music
Drawing and handiworks
Home Affairs
Sports
Morals
and special activities

In private schools religion may be added or exchanged with morals.

The number of hours and their allocation are also set in the Law.

Teaching of any subject, including statistics, has to find its place somewhere in such a scheme.

It was one of the important issues where to include teaching of statistics. Statistics are used and taught in sciences and social discipline but the basic and systematic, if one may say so, lectures were decided to be given in mathematics.

The goal of the educational course for mathematics states, "To provide the basic knowledge and technical abilities pertaining to quantity and figures and to build up the ability and attitude which enables to understand mathematically, think systematically and reasonably and handle the events occurring commonly." Teaching statistics is also included in this goal.

The goal is further sub-specified with relevant contents for each grade. For such purposes the field of educational course of mathematics is sub-divided into:

A. Numbers and computation
B. Quantities and measurement
C. Figures
and
D. Quantitative relations

and statistics is understood to belong to D.

The present Guideline was revised in 1977 superseding the one issued in 1968. At that time "to provide less tight but complete school life" was one of the principles, and it was suggested, in short, to reduce the contents, limiting only to the basic and those to be repeatedly treated are to be simplified.

Thus, if we compare the two Guidelines, we note the teaching of statistics is also simplified or deferred in general.

In the current Guideline, explicit sub-division of D (Quantitative Relations) into "presentation by formula" and "statistics" has vanished. Statistical subjects start from the third grade while they were from the second in the previous one. The term "probability" was included in the former one, but it is deleted from the present. The parts of the Guideline where statistics is explicitly or implicitly referred to are as follows:
"To arrange in order", "to sort and count", "to express in the form of tables and/or charts", and "to read tables and/or charts" appear here and there but not necessarily in connection with statistics. Thus, such were not included in the following:

Grade 3

Goal No.4
Through arranging the material and using formulae and graphs, improve the ability to express and study quantities and their relationship.
(One of the 4 goals for Grade 3).
Contents
D. Quantitative relationship
(2) To be able to express the data in, and read them from, tables and/or charts.
  a. To sort and/or arrange the data in tables on the basis of time and place.
  b. To learn how to read and draw bar graphs.

Grade 4

Goal No.4
To improve the ability to express or study quantities and/or their relationship on the basis of formulae and graphs as also to be able to investigate dependency and/or classify and arrange according to the objective.

(one out of 4)
Contents
B. Quantitative relationship
(1) To improve the ability to express or study the relationship between two related variates;
a. For simple cases, to study the relationship through considering the corresponding quantities and/or expressing the set of values in the form of a table,
b. To express the change in line graphs and/or to understand the characteristic nature of the change from such,
(3) To improve the ability to collect, sort and arrange and study the characteristic feature according to the purpose,
a. To study the cases of occurrence for two events.
b. To check the omission and duplication of data.
c. To express the data in bar graphs and line graphs or to study the characteristic feature and trend from graphs.

Grade 5
Goal No.4
... Also, to improve the ability to study statistical data through the use of percentages and pie graphs.
(one out of 4)
Contents
B. Quantity and measurement
(3) Deepen understanding of the observation of the magnitude of quantity and the meaning of the observed values;
...
(b) To understand the meaning of the average and use it;
B. Quantitative relationship
(1) Explain the meaning of percentages and enable the pupil to use them;
(4) To be able, according to the purpose, to sort and arrange the data, express them in pie graphs and bar graphs;

Grade 6
Goal No.4
Along with improving the ability through the understanding of proportion etc. to consider functionally, improve the ability to consider and/or express statistically, e.g. studying the distribution of the data.
Contents
A. Numbers and computation
(2)
e. To understand that numbers can be used to express the degree of occurrence of undeterministic events.

D. Quantitative relationship

(3) To improve further the ability to consider and/or express statistically, - e.g., to study the dispersion of the data for simpler cases.

a. To learn about tables and/or charts expressing the frequency distribution.

b. To learn there are cases where the tendency of the entirety can be known by, for example, rates obtained from partial data.

c. To select adequately tables and/or charts according to the objective and to devise something convenient.

(4) To develop gradually the ability to study the possible cases for occurrence through the orderly arranging of simple events.

The annotated part of the Guideline provides goals as follows:

Middle grades

> ... to improve the ability to express and consider quantities and their relationships with formulae, tables, charts and the like as also to improve gradually the ability to study the dependency, characteristic feature of, through sorting and arranging, the materials, according to the purpose. ...

Higher grades

> ... main objectives are ... to improve the ability to consider statistically, e.g. to survey the distribution of data ...

To achieve the goals set the contents of the syllabuses are grouped into the 4 regions of

A. Numbers and computation
B. Quantities and measurement
C. Figures
D. Quantitative relationship

D is a little different from others and may be further grouped into, (a) understanding of function, (b) expressing and/or interpreting formulae, (c) statistical consideration, (d) understanding of aggregation, etc. ...

D is not provided for grades 1 and 2.

c) Statistical consideration

The objectives here are to find out the relationship, tendency and rules existing, through observing as an aggregate, out of events which are governed by randomness, each element of which does not seem to follow any rules.

(1) Characteristic feature of the studies

In order to improve the attitude and ability to consider statistically, let the pupil study a series of processes characterized as follows so that they master the knowledge, techniques and fundamental thoughts involved in them:

a) Clarify the problem and objectives and collect appropriate materials. Then quantify them adequately.

b) Sort the data from the viewpoint based on the objective and facilitate grasping of characteristic features.

c) Interpret what kind of characteristic feature or tendency the aggregate has, through sorted data and/or means and ratios.

d) Infer on the basis of partial materials, what kind of statement can be made with regard to the population. Further engage them in inferring the tendency and/or characteristic feature not covered by the materials, and trying to solve the problems.

In the middle grades b) and c) are highlighted, to provide the basic knowledge and techniques of statistics.

Starting with grade 4 such contents as to collect materials on the basis of the purposes, to consider the general tendency ... are to be given.

(2) Cases for lower grade

As given, for example, "to sort simple cases and express them in tables and/or charts..." cf. (1) for grade 2, it is necessary to give basic experiences of statistical processing of sorting and then facilitating understanding of the characteristic feature of the data and make accustomed with such viewpoints.

Contents for respective grade

Grade 3

- Arranging data and graphs
  - a) Sorting and arranging data
  - b) Reading and drawing bar graphs

Grade 4

- (11) Relationship between two dependent variables
  - a. Detecting rules
  - b. Reading and drawing line graphs
- (13) Sorting and arranging data and graphs
  - a. Cases of occurrence of two events
  - b. Study of characteristic feature and/or tendencies

Grade 5

(11) percentages

(12) Relationship of quantities expressed by a simple formula

(14) Pie graphs, belt graphs

Grade 6

(2)

c) Expressing by a number, the degree of occurrence of undeterministic events

(8) Dispersion of data

a) Tables and/or charts for frequency distribution

b) General tendency

(9) Possible cases of occurrence

Complete sets of textbooks for the educational course in mathematics are to be shown with some explanation in the meeting.

The author feels the basic skeleton of the framework of statistics education at elementary education has been provided, although there remain many other aspects needing clarification, e.g. pertaining to teachers, their qualification, special courses, additional materials to mention just one genre. Naturally to present a comprehensive view of statistics education in elementary education is too gigantic a problem and cannot be done here. Thus,the author intends to present one more aspect only.

The Japan Statistical Society is the professional institution representing the field of statistics with about 1,000 members. Some members are interested in the matter of statistical education in elementary education but they are a very minor group and reports pertaining to such do not appear significantly. Besides JSS, there are many professional societies in the field of statistics, e.g. Mathematical Society of Japan with a section on Mathematical Statistics, Applied Statistics Society, Research Association of Statistical Sciences, etc. The situations are the same and very few or no teachers in primary or lower education have very limited contacts with such as they are organized in other associations and societies.

The Society for Mathematical Education in Japan and a few others are for mathematical education, thus concerned with statistics education, as explained elsewhere.

However, a unique association of statistics education which cannot be missed is the All Japan Council of Statistical Education Research as the association of prefectural Council of Statistical Education Research established in nearly all prefectures with memberships of 1,000-2,000 each.

Each of the prefectural Councils organised the training sessions and the research meetings every year and the national ones are conducted with about 1,500-

2,000 participants. They issue biannually "Statistics Education Research" incorporating reports of those sessions.

They have issued "Manual for Statistical Education", which is utilized extensively. The manual has been quite influential for statistical education in primary education.

The Council issued in 1980 the first collection of the Actual Examples of Statistical Education, which provides good examples of actual statistics education in primary schools.

TOPIC 2: TEACHING STATISTICS IN SCHOOLS

(12-15 YEARS AGE GROUP)

Organiser: Professor P L Hennequin

WHY TEACH DESCRIPTIVE STATISTICS IN SECONDARY SCHOOLS?

Heinrich Winter, Rheinisch-Westfälische Technische Hochschule

"A man standing with his one foot on a hot stove and on ice with the other enjoys - in terms of statistics - pleasantly warm feet. That is the logic of the average". (H. Paulsen - a German doctor - in: DER SPIEGEL, Number 43, 1978, p.217).

In this paper I intend to propose some arguments in favour of Descriptive Statistics (DS) in the school curriculum. I think every pupil has to learn some elementary procedures, skills and concepts concerning gathering, representation, summarisation, characterising, analysing and interpretation of data. The main aim of this paper is to convince teachers in Secondary Schools both of the importance and teachability of DS; therefore, I will give some practical suggestions too. Let me present my paper in the form of 5 theses; the first 4 may be called descriptive and the 5th prescriptive.

THESIS I: DS is more substantial than many people assume.

(a) DS is more than computing averages or standard deviations of lists of numbers. Rather, DS concerns describing and modelling reality in which chance plays an important role as to quantities and values of any real entities undergoing variations and fluctuations.

Many pupils fail, when they have to solve a problem of common "deterministic" real life mathematics, even if it is easy with regard to computation. The main difficulty consists in mapping a real life situation onto an appropriate mathematical model. If the real life situation is indeterministic, then its mathematical modelling requires much more mental effort. The reasons are: (1) The regularities or laws to be seen in the situation are often quite hypothetical and may come into conflict with intuitive ideas. (2) The data must often be gathered by the pupils themselves and appear variable and unsteady. (3) The representation and processing of the data are not carried out by simple and fixed rules. (4) The interpretation of the solution requires a feeling for the special factors involved in the problem.

In this respect, DS can be considered as too demanding rather than as too trivial.

Let us have a look at some illustrative examples.

(i) Average temperature of a day

How can we define "the" average temperature of a day? Who is interested in this information? How exact should this average value be?

Pupils might suggest:

$$x_A = \frac{x_{max} + x_{min}}{2}$$

This definition is indeed common (e.g. in the USA); and the average is easy to find using a maximum-minimum thermometer.

But discuss this definition. Construct cases in which this average is misleading, is not misleading.

Assume, the "real" average (as arithmetic mean) is $\bar{x}$. Estimate the difference between x_A and x. It can be shown:

$$|\bar{x} - x_A| \leq \frac{x_{max} - x_{min}}{2} \quad \text{(a little theorem!)}$$

Proof:

$$x_{min} \leq \bar{x} \leq x_{max}$$

$$x_{min} - x_A \leq \bar{x} - x_A \leq x_{max} - x_A$$

$$-\frac{x_{max} - x_{min}}{2} \leq \bar{x} - x_A \leq \frac{x_{max} - x_{min}}{2}$$

$$|\bar{x} - x_A| \leq \frac{x_{max} - x_{min}}{2}$$

Another definition (also used by geographers) of average temperature is:

$$x_B = \frac{x_1 + x_7 + x_{13} + x_{19}}{4} \qquad x_i \text{ means: Temperature at i o'clock}$$

A new problem: Compare $\bar{x}$, x_A, x_B.

Assume the temperature-curve of a day is given on graph paper. How can you find an average temperature from that? In fact: We have actually a problem of calculus! But we can use the temperature-curve for computing an average temperature in different ways. We can easily add similar problems: average temperature of a month, of a year, and so on.

(ii) School report

A teacher (in Germany) has to judge his pupils' achievements in the last half year. He has to find a final mark (from 1 to 6. 1 = very good, ..., 6 = poor). The marks from 6 tests are the basis of the final decision. Here are the test marks of two pupils, A and B.

Test Number	1	2	3	4	5	6
Test marks of pupil A	2	3	3	1	3	6
Test marks of pupil B	4	4	4	4	1	1

Can you propose a fair final report mark for each pupil? Is A better than B? Why? Why not? Under which circumstances is B better than A?

Here are several possibilities of combining the 6 single marks.

Final mark by	Report mark of A	 of B	Result
Arithmetic mean	3	3	A as good as B
Median	3	4	A better than B
Modal	3	4	A better than B
Maximum	6	4	B better than A
Minimum	1	1	A as good as B
33 1/3 % value	2	1	B better than A
66 2/3 % value	3	4	A better than B
Mid range	5	3	B more constant than A
Variance	2 1/3	2	
Function in time	oscillates	decreases	B becomes more and more better

Well, what is "really" a fair report mark? Each judgement may be wrong. But we must finally come to a decision. The statistical processing of the data helps us only to describe our decision; we need, however, extra-mathematical arguments to justify our decision.

(iii) Pocket money

Someone may say that the higher the children's pocket money the higher is the income of their parents.

Is that a true statement? How can we test it?

A classroom experiment may be carried out - in fact a delicate matter - perhaps by means of an anonymous interview. Each pupil writes two numbers on a slip of paper:

My pocket money ..... DM
My parents' monthly income..... DM

We collect the slips, mix and numerate them. We note the information in tabular form, for instance (some fictitious figures!):

Nr.	x	y	Nr.	x	y
1	20	2 000	11	20	2 500
2	20	2 300	12	15	2 400
3	15	2 200	13	15	2 000
4	10	1 500	14	10	2 000
5	-	2 400	15	40	2 900
6	15	2 100	16	30	3 200
7	20	1 800	17	10	1 800
8	25	1 500	18	-	1 400
9	30	3 000	19	30	2 100
10	25	3 400	20	30	2 200

x: pocket money in DM; y: monthly income in DM

This table shows the data in a badly arranged way. How can we make it more transparent?

Here is a lucid possibility, the so-called stem-and-leaf display:

pocket money (DM)

0	00
1	0005555
2	000055
3	0000
4	0

incomes (100 DM)

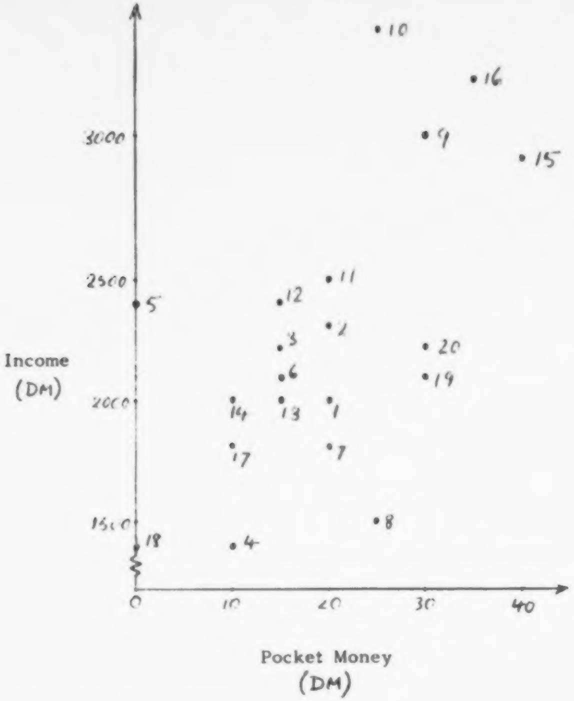
1	45588
2	000112234459
3	024

Obviously this is a better presentation than the traditional histogram. We have a "numerical sight" of the frequency distribution, too.

We are, however, mainly interested in the relation between the two random variables: pocket money, income. Therefore, we need a simultaneous representation.

A scattergram gives us a first answer about how allowance is correlated with income (in our sample!).

At the first glance the scattergram (next page) presents a confused picture; it seems to be a muddle. But on closer examination we may discover a trend which corresponds to our statement: the higher the pocket money the higher the income. On the other hand, we see many "exceptions", too. Can we calculate the extent of the correlation in some way?



A four fold table might be suggested:

y>2100	3,5,12 d=3	2,9,10,11,15,16,20 b=7
y≤2100	4,7,13,14,17,18 a=6	1,7,8,19 c=4
income y x allowance	x ≤ 15	x > 15

Pupils may discover a measure of association, such as:

$$k = \frac{(a + b) - (c + d)}{a + b + c + d}$$

In our case, k is (13 - 7) : 20 = 0.3.

Now discuss k.

As is well known, the usual measure of correlation is:

$$f = \frac{ab - cd}{((a + c)(b + d)(a + d)(c + b))^{1/2}}$$

but it is surely impossible to explain it to pupils of this age.

Pupils should be encouraged to find another approach to our correlation problem.

We note from our sample the rank of each pupil according to his allowance and to his parents' income. We obtain this new table:

Number of pupil	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Allowance rank	10	11	6	3	1	7	12	14	16	15	13	8	9	4	20	17	5	2	18	19
Income rank	6	13	11	2	14	9	4	3	18	20	16	15	7	8	17	19	5	1	10	12

If the allowance would be proportional to the income then the lower rows would be identical. But only the 17th pupil has the same rank in both cases. An obvious measure of association might be suggested: the arithmetic mean of the absolute differences between the respective rank numbers.

In our case we obtain 90/20 = 4.5. But this is a bad measure, because it is not standardized.

Let's seek a better measure. Look at the rank pairs $\binom{10}{6}, \binom{11}{13}, \binom{6}{11}$

Compare all combinations of rank pairs; there are 190 cases. We can define:

$$\binom{x_a}{x_i} \text{ is concordant } \binom{y_a}{y_i} \quad \text{if} \begin{cases} x_a < y_a & \text{and} \quad x_i < y_i \\ & \text{or} \\ x_a > y_a & \text{and} \quad x_i > y_i \end{cases}$$

$$\binom{x_a}{x_i} \text{ is discordant } \binom{y_a}{y_i} \quad \text{if} \begin{cases} x_a < y_a & \text{and} \quad x_i > y_i \\ & \text{or} \\ x_a > y_a & \text{and} \quad x_i < y_i \end{cases}$$

For example: $\binom{1}{14}$ is concordant $\binom{15}{20}$; $\binom{1}{14}$ is discordant $\binom{7}{9}$.

Now we classify all 190 possible cases. If P is the number of concordant, Q the number of discordant cases, then

$$g = \frac{P - Q}{P + Q}$$

is a reasonable measure of correlation. Here we find $g = \frac{130-60}{190} \approx 0.368$.

(b) DS as a part of stochastics probably does not comprise as many ideas, problems, theorems as e.g. geometry does; but one should not consider DS merely as a collection of poor tools applied in non-mathematical fields. On the contrary, there are some mathematically interesting concepts and theorems. For instance, we may discover, formulate and prove the following theorems:

(i) $\overline{(x + a)} = \bar{x} + a$

(ii) $\overline{(x \cdot a)} = \bar{x} \cdot a$

Arithmetic mean as a linear concept

(iii) $\bar{x}$ is the centre of gravity of the frequency distribution.

(iv) $V = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 = \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2$

(v) Minimum property of the mean $\bar{x}$ with respect to the variance

$$V = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x})^2 \leq \frac{1}{n} \sum_{i=1}^{n} (x_i - x)^2 \quad \text{for all } x.$$

(vi) Minimum property of the median $\tilde{x}$ with respect to the arithmetic mean of the absolute differences:

$$D = \frac{1}{n} \sum_{i=1} |x_i - \tilde{x}| \leq \frac{1}{n} \sum_{i=1}^{n} |x_i - x| \quad \text{for all } x.$$

and so on.

Proof for (v) without using any methods of calculus:

$$\frac{1}{n} \sum_{i=1}^{n} (x_i - x)^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i^2 - 2x_i x + x^2)$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2\, x \cdot \frac{1}{n} \sum_{i=1}^{n} x_i + \frac{n}{n} x^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 - 2\, x\, \bar{x} + x^2 + \bar{x}^2 - \bar{x}^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} x_i^2 + \underbrace{(\bar{x} - x)^2}_{\geq 0} - \bar{x}^2$$

$$\geq \frac{1}{n} \sum_{i=1}^{n} x_i^2 - \bar{x}^2 = V \text{ (iv)}$$

These few statements might have sufficiently supported my thesis 1, namely DS is more substantial than is generally assumed.

THESIS 2: Teaching DS fosters general aims of education in compulsory schools

(a) If school education aims at the development of children's abilities to communicate in their daily and public life, then teaching DS is helpful. In particular, pupils can and must learn to understand, to criticize and to apply numerical or half-numerical arguments as they are widely used in our society.

Among other things, pupils should be enabled

(i) to read the information contained in graphs, diagrams and tables which we find in newspapers, in books, on television, ...,

(ii) to draw instructive graphs,

(iii) to detect misuses of (iconic) representations of data,

(iv) to become aware of the meaning of half-numerical expressions in natural language, like: roughly, approximately, about, round, almost, at least, never, seldom, by chance, ...,

(v) to distinguish between a numerical statement itself and its biased presentation, e.g. "What a disaster: Prices have rocketed up by 10% since last year!" - compared with: "Things are getting better: only 10% price rise since last year!",

(vi) to insist on further or possibly on complete information (for example, upon the statement "Homework of a 7th degree pupil lasts 2 hours a day" it is necessary to ask "Is 2 hours the minimum, maximum, median, modal, the arithmetic mean, or whatever? How was this value found, by a random sample? How large was the sample?"),

(vii) to unmask specious numerical arguments (for example: "In such and such a state house building is increasing greatly: This year 50% more houses were built than in the preceding year." - "Mother's milk is very healthy, for 9 in 10 of 90 years old people were breast-fed in their infancy."

(b) If school education aims at supporting the pupils' abilities to understand their environment, then we should teach DS. There are, around us, a lot of indeterministic processes. Pupils should become aware of, and, at least, basically understand randomness in our daily world. I give some examples: Weather in our country, weight of eggs, height of 7-year-old boys, lifetime of people in our country, quality of an industrial product, service times (at filling stations, telephone boxes, in shops,......), nominal values and problems of calibration, prices and their ups and downs, peoples' consumer behaviour, illnesses and medicines, events in sport, and so on.

(c) Statistical methods of various levels infiltrate steadily into many jobs, even into the domain of handicraft. Random tests are increasingly used in order to control the quality of mass products, to control the nominal values (e.g. the weight of packets of sugar), to prove the effectiveness of a medicine, estimate the number of future customers of a product, to assess damage caused by hail, and so on. Therefore, DS should be taught, as school education has to prepare children for their future professional life.

THESIS 3: The methods of, and the pursuits in, DS may concern, probably more than other mathematical activites, the pupil's actual existence.

In my view, this thesis presents the prominent reason for teaching DS in compulsory schools.

We often notice a lack of motivation for learning in classrooms, in particular in Secondary School classrooms, whatever the reasons are for that. Pupils are (openly or covertly) uninterested in most learning objectives. They obviously believe that school learning does not solve any problems of their actual life. I, of course, cannot offer suggestions, which would surely improve this deplorable situation. I think, however, DS gives a good chance to foster motivation for learning, at least for some pupils.

(a) Examples can easily be found in pupil's actual surroundings. There are a lot of subjects attracting his interest: sports, games, leisure time; school life; motor vehicles, traffic; illness and death; election forecast; and so on.

It is important to engage pupils in the choice of the subjects and in the collection of the data. The data should generally not be fictitious but as realistic as possible. The more authentic the problems are the stronger a pupil's interest and activities will be.

(b) Learning DS is learning important aspects of human life. Pupils become aware of life being an uncertain matter. Preferred subjects should be those which stimulate a pupil's thinking about his own existence now and in future.

A pupil can learn to realize that nobody and no institution will be able to guarantee a long life, a healthy life, an economically successful life, a satisfying job, to find a trustworthy partner, and so on. The future is dark and unknown, life is full of risks. A pupil, nevertheless, can learn how to provide for the future. Thus, every pupil should e.g. obtain a basic knowledge of insurance. This matter being rather complicated should be simplified in a responsible way.

Children also become aware of certain aspects of human learning and behaviour. Results of restricted observations are often generalized and decisions are made on grounds of observations. For instance: "A number of times I have bought a book in a particular book shop, and each time the shop assistant was friendly and knowledgeable. So, I think this is an excellent book shop". In the classroom might be discussed: How large does the number of observations have to be?

Can we learn on the basis of only one observation? What actually do we do in generalizing results from limited observations? Is the generalization grounded in the observations? Yes, but isn't it more than that? Do we not superimpose some mental construction on the data: do we not induce an hypothesis? How sure is such an hypothesis? How can we corroborate it? Can we interpret the data by means of the hypothesis? What should we do when new observations do not fit the hypothesis? In particular, children can learn to criticise the naive use of common sense and, above all, the role of prejudices.

Another aspect of human behaviour should be led to awareness, I mean actions on grounds of intuitively assigned probabilities. E.g., we switch on the light, but the bulb stays dark. Why? We suppose: very probably the bulb filament is broken, or the fuse has gone. We hardly suppose: the power station has been blown up, although this would be possible. Thus, we would change the bulb rather than collect some money for a new power station. Another example: we go to the bus stop, although all the bus drivers may have fallen ill. This is reasonable because the probability of some epidemic among bus drivers is rather small. Obviously, we assign probabilities to the various events of daily life - rough probabilities, but mostly precise enough for common actions; and these probabilities originate from former experiences.

(c) By learning DS pupils experience direct applicability of mathematics. The pleasant feature of DS - unlike mathematical fields - is that pupils are concerned with interesting and significant problems without being obliged to use complicated and technical concepts and methods.

Some concepts of DS can immediately be applied. For instance: The modal of a sample is indeed a simple notion, and is extremely easy to find in the value table or in the frequency distribution of a (random) sample. But how significant are the applications of the modal in real life situations! Politics: absolute, simple, bare, vast majority in parliaments; majority votes; coalitions; plural voting; - Society: pressure by majority opinion; problems of minorities; influence of public opinions; - Checkings: Is the most frequent result always the "true" result? - Children can often directly watch statistical activities in their real world: customs examination; traffic census; traffic control; investigation of air or water or soil, and so on.

(d) Finally DS might offer the chance for one or other pupil to try a new start in learning mathematics at all. The immediate applicability and the small effort of conceptual developments could be a basis for motivating those pupils who have been unsuccessful in mathematics or have hated mathematics or have suffered from anxiety.

Perhaps the teacher attempts a new start and his work becomes more interesting and more successful.

THESIS 4: DS can easily be merged into the canonical syllabus without overloading it.

(a) DS should not appear well separated from classical school topics. On the contrary, DS can be closely interwoven with other common topics of mathematics. In addition, DS projects can be chosen from every other school subject.

Mathematical topics:

Set theory: sets of statistical units; samples as subsets; frequency distributions as functions; classifying; ranging.
Fractions: relative frequencies; percentages.
Natural numbers: numerations; combinatorics.
Geometry: graphs, Monte-Carlo Methods.
Square equitations: variance; standard deviation.

Other subjects (examples):

Geography: weather; trade; production; soils.
Biology: life spans, growth; sizes.
Physics/Chemistry: measurements.
Road Safety Education: accidents.
Language: texts.
Economics: prices; incomes; production.

A further advantage of DS is that most of the topics can be described only by means of natural language, so that there is no need for endless exercises to make the pupils familiar with elaborated formula notations.

(b) Thus, there is little danger of overloading the canonical syllabus; DS, instead, may enrich aspects of other topics in mathematics.

Let me illustrate this through two examples:

(i) Fractions
Many children consider fractions like $\frac{29}{134}$ as monsters lacking in clarity and practical significance. The DS approach to fractions (≤ 1) makes more sense: $\frac{29}{134}$ should be interpreted as a relative frequency, for instance: Among 134 interviewed persons 29 preferred TV programme A. Is that a big proportion? Do people in general like this programme A? The situation leads to the transformation of $\frac{29}{134}$ into decimals or percent, 0.216 ... 21.6%, and this may be interpreted: About 22 in 100 persons preferred programme A, that is about 1 in 5 persons.

DS also offers further motivation for computing fractions, even multiplying fractions (≤ 1) : If f(a) is the relative frequency of the attribute a, f(b) the relative frequency of the attribute b, and if a is independent of b, then f(a) f(b) is the relative frequency of the combined attribute a and b. Now it also follows if two fractions are both smaller than 1, then the product is smaller than each of the fractions.

(ii) Functions

The notion of function is (at least since Felix Klein) a main and guiding idea, inspiring mathematics throughout.

DS supplies a host of various examples of functions:

- individual $x \longmapsto$ interesting attribute of x
- attribute $a \longmapsto$ relative frequency of a
- sample $S \longmapsto$ mean value of the sample S

and so on.

More important is the fact that DS may bring forth a deeper understanding of what a function essentially is.

For instance, we gather the data on distances and the respective fuel consumptions of a certain car. The driver noted two values each time he filled up his fuel tank. He may have obtained this table:

distances (km)	fuel consumption (l)
325	35.8
340	38.1
309	33.4
182	20.1
405	43.9
265	29.1
352	38.7
315	34.6
340	37.5
352	38.1

It seems clear that there is a strong dependence: the longer the distance, the more petrol was consumed. But is that exactly true? Does the table represent a function in the mathematical sense? Is there only one single value of fuel consumption for a given distance? Obviously not for

$$\begin{tikzcd}[row sep=tiny] 340 \arrow[r] \arrow[dr] & 38.1 \\ & 37.5 \end{tikzcd}$$

It is not unequivocal on the right. If any distance out of our table is given, then we can not always find only one corresponding consumption. Therefore, the table above is not the set of pairs of a function. On the other hand, we are not far from an exact function. We can write down a simple function approximately fitting the data:

$$\begin{aligned} f:\ \text{distance (km)} &\longrightarrow \text{fuel consumption (l)} \\ x &\longmapsto 0.11\,x = f(x) \end{aligned}$$

However, is this function the best approximation? We found it by using our eyes! Is there a method to calculate best approximations? Well, we are very close to regression problems.

THESIS 5: In order to benefit by the educational value of DS, mathematical modelling of problems in the pupils' surroundings should be one of the main and fundamental activities.

This is a prescriptive thesis in so far as it contains some suggestions for classroom activities.

Teaching of DS will not produce any relevant educational effect, if we do not explicitly concentrate our efforts on the goals mentioned in the previous theses 1 to 4. Teaching DS will not automatically be successful in the sense I was trying to describe. Of course, mechanical skill training and the repeated solving of small tasks, all of the same type, is totally inadequate.

(a) At the start, one should have a challenging real life situation stimulating the pupils to talk and to think, to ask questions and to discover problems.

The extra-mathematical aspects of the situation must be carefully respected; these aspects should not be misused as merely methodical decorations. It is strongly recommended that pupils carry out projects by interaction with other branches of study.

(b) Pupils must become aware both of the powers and of the limits of statistical procedures. In particular, children should learn that most of the data are blind, useless and unsuitable for drawing any conclusions or for establishing any rules.

(c) At any rate, DS must not be reduced to description in a too strict sense. DS should be regarded as a part of stochastics. Therefore, aspects of probability must be incorporated.

(d) DS should be taught according to the principle of curriculum spiral. The notions and procedures of DS must gradually grow and be related to other mathematical ideas. DS should not be an isolated and separated subject, but should rather be an instructional principle.

To illustrate the 5th thesis let me outline an example:

Life and Death

(i) Possible starting points

- a TV discussion about problems of old age insurance (decreasing working population, increasing number of pensioners)
- a fatal traffic accident in our street
- a newspaper report about infant mortality in developing countries

The classroom discussion could bring up the question:
What is the life span of man?

(ii) First explorations

We seek information.

We look up the obituary notices in a newspaper.

For instance we find in the "Aachener Volkszeitung" (February 9th 1980; authentic figures!):

Age (of dying)	83	70	69	75	95	71	80	77	43	82	85	65	71
Sex	f	f	m	m	f	m	m	f	m	f	m	m	m

What does this table show? Try to concentrate the data. Do you dare to make any speculations?

We find averages:

Arithmetic means			Medians		
	74.3	(total)		75	(total)
	69.9	(male)		71	(male)
	81.4	(female)		82	(female)

It is worth noticing: Most people had a life span longer than 70 years. On average, women had a longer life than men.

Can we, however, generalise this observation? How reliable is our source?

Further sources of data are:

- cemetery
- municipal administration
- official statistical books

(iii) Discussing a mortality table

The teacher presents the mortality table of the next page. The table might stimulate to discuss and solve problems like these:

- How can we obtain a good survey of these 100 numbers?

  Draw a cartesian graph of the function

  $$\begin{array}{lcl} \text{age } x & \longmapsto & \text{living men of age } x \\ x & \longmapsto & l(x) \end{array}$$

  What strikes you? Describe this "mortality curve".

- Investigate the mortality per year. Where is the minimum, the maximum? What is particularly striking?

  Assume, the mortality per year would be constant!

  We define:

  Probability of death at age $x = p(x) = \dfrac{l(x) - l(x+1)}{l(x)}$

  This is the probability of dying in the course of the next year if someone is x years old.

  Draw a cartesian graph of the new function:

  $x \longmapsto p(x)$

General Mortality Table 1970/72 Germany (FRG) Male

x	l_x	x	l_x	x	l_x
0	100 000	35	93 245	70	54 909
1	97 400	36	93 049	71	51 838
2	97 249	37	92 838	72	48 673
3	97 152	38	92 610	73	45 438
4	97 067	39	92 361	74	42 161
5	96 989	40	92 089	75	38 872
6	96 918	41	91 794	76	35 601
7	96 854	42	91 475	77	32 373
8	96 795	43	91 131	78	29 212
9	96 741	44	90 761	79	26 137
10	96 692	45	90 363	80	23 167
11	96 647	46	89 934	81	20 321
12	96 604	47	89 468	82	17 619
13	96 561	48	88 958	83	15 083
14	96 515	49	88 398	84	12 735
15	96 459	50	87 787	85	10 595
16	96 383	51	87 104	86	8 678
17	96 273	52	86 369	87	6 990
18	96 118	53	85 574	88	5 529
19	95 927	54	84 717	89	4 287
20	95 732	55	83 789	90	3 253
21	95 541	56	82 779	91	2 407
22	95 357	57	81 673	92	1 735
23	95 182	58	80 460	93	1 213
24	95 016	59	79 130	94	824
25	94 858	60	77 675	95	539
26	94 705	61	76 087	96	339
27	94 555	62	74 357	97	204
28	94 405	63	72 477	98	117
29	94 253	64	70 440	99	64
30	94 097	65	68 242	100	33
31	93 937	66	65 882		
32	93 773	67	63 361		
33	93 604	68	60 685		
34	93 429	69	57 864		

x: age

l_x: living men of age x

Source: Statistiches Bundesamt, Bevölkerung und Kultur, Kohlhammer, Stuttgart 1976

Describe the probability of surviving: $x \longmapsto 1 - p(x)$

- A man is 40 years old today. What is the probability that he can celebrate his 65th birthday?
- What is the probability that all 3 new born babies are still alive in 50 years?
- Compute some probabilities concerning yourself.
- Can you find the median and the arithmetic mean of life span by using the table?

It is easy to find the median (roughly 71 years). To find the arithmetic mean, however, demands some calculations. It is useful and not very difficult to display the formula:

$$\bar{x} = 0.5 + \left(\sum_{x=1}^{100} l(x) \right) : 100\,000$$

Here is $\bar{x} = 67.4$. This value is also called "average life expectancy at birth". Find the remaining average expectancy of 15 (30,50,77,......) years old men.

- It is a quite expensive matter to make a mortality table. How is it made? Who makes it? The teacher should outline how such a table is set up, emphasising the basic ideas.

- Who might be interested in mortality tables?
  In a simple but responsible way the fundamental ideas of life insurance should be developed and some concrete examples should be considered. We calculate (e.g.) the life insurance premium rate a man has to pay under the following simplified conditions: given his age of insurance beginning, the amount of the policy, the length of the endowment period, the interest rate. The company has to pay the amount either at death or at the end of the endowment period, whichever occurs first. The premium has to be payed at once.

(iv) Comparing further mortality tables.

Excerpts from Mortality Tables from 1871/80 to 1970/72 (Germany):

	MALE						FEMALE					
AGE	1871/80	1910/11	1932/34	1949/51	1960/62	1970/72	1871/80	1910/11	1932/34	1949/51	1960/62	1970/72
0	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000	100 000
5	64 871	77 213	89 654	92 880	95 629	96 989	68 126	80 077	91 535	94 270	96 782	97 690
10	62 089	75 984	88 793	92 444	95 620	96 692	65 237	78 816	90 753	93 937	96 579	97 492
15	60 892	75 189	88 244	92 097	95 388	96 450	63 678	77 930	90 270	93 701	96 434	97 399
20	59 287	73 832	87 298	91 466	94 812	95 732	62 324	76 659	89 490	93 295	96 188	97 059
25	56 892	72 130	86 032	90 531	93 940	94 858	60 174	75 043	88 390	92 711	95 884	96 756
30	54 454	70 425	84 715	89 518	93 166	94 097	57 566	73 115	87 139	92 039	95 485	96 429
35	51 815	68 545	83 234	88 428	92 322	93 245	54 685	71 020	85 754	91 221	94 940	95 997
40	48 775	66 227	81 481	87 102	91 218	92 089	51 576	68 659	84 135	90 225	94 184	95 331
45	45 272	63 238	79 285	85 342	89 659	90 363	48 481	66 187	82 211	88 901	93 081	94 308
50	41 228	59 349	76 322	82 640	87 230	87 781	45 245	63 231	79 620	86 901	91 442	92 683
55	36 544	54 290	72 147	78 562	83 221	83 789	41 308	59 390	76 038	84 225	89 063	90 272
60	31 124	47 736	66 293	72 852	76 652	77 675	36 293	54 016	70 984	80 166	85 404	86 903
65	24 802	39 527	58 106	64 999	66 941	68 242	29 703	46 404	63 712	73 875	79 839	81 647
70	17 750	29 905	47 059	54 394	54 461	54 909	21 901	36 440	53 104	63 994	70 840	73 157
75	10 743	19 328	33 479	40 700	39 784	38 872	13 677	24 517	39 132	49 605	57 076	60 033
80	5 035	9 711	19 122	25 106	24 156	23 167	6 570	12 981	23 500	31 787	38 507	42 046
85	1 635	3 297	7 732	11 321	10 861	10 595	2 232	4 794	10 323	15 225	19 500	22 478
90	330	679	1 966	3 176	3 082	3 251	471	1 126	2 868	4 815	6 480	8 016

Source: Statistisches Bundesamt, Bevölkerung und Kultur, Kohlhammer,
Stuttgart 1976

What strikes you?

- Investigate the development of infant mortality.
- Investigate sex differences. What do you think about that?
- Find the approximate medians and the arithmetic means for the different times.
- Make inquiries about mortality in several other countries.

References

[1] Barnett, V. (1980) Teaching Statistics in Schools in England and Wales in: Comparative Studies of Mathematics Curricula, Change and Stability, Institut für Didaktik der Mathematik, Materialien und Studien, Band 19, Bielefeld, p.444-463

[2] Dinges, H. (1982) Deskripive Statistik für die Mittelstufe, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.5-27

[3] Engel, A. (1982) Statistik auf der Schule: Ideen und Beispiele aus neuerer Zeit, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.27-85

[4] Ganadesika / Kettering/Tukey (1982) Themen aus der Datenanalyse: Begriffe, Beispiele und Pädogogik, in: Der Mathematikunterricht, 28, Klett Stuttgart, p.28-56

[5] Heller, W.D. (et al) (1979) Beschreibende Statistik, Birkhäuser Basel

[6] Winter, H. (1981) Zur Beschreibenden Statistik in der Sekundarstufe 1, in: Dörfler/Fischer (Ed.): Stochastik im Schulunterricht, Hölder-Pichler-Tempsky-Teubner Wien-Stuttgart, p.279-303

THE TEACHING OF STATISTICS IN FRENCH SECONDARY SCHOOLS

G Dumousseau, IREM de Bordeaux

Statistics is not at present taught to French 11 to 15 year olds. Syllabuses do not require it and text books offer very few activities or exercises bearing on statistics. Very simple statistical sequences may be found in some books, in the sections on proportionality, percentages or reading graphs.

Mathematics masters, who generally are not trained in statistics and who are not encouraged officially to use this subject, practically ignore it. History and geography masters employ more statistics, albeit in simple examples, like reading a chronological sequence in a table, or following the evolution of some economic variable for which the statistical treatment has already been carried out.

There are, however, many teachers who feel it is unreasonable that children should be left solely in the deterministic framework, and essential that they should also have contact, as soon as possible, with numerical data. Over the last decade, the research institutes for the teaching of mathematics (I.R.E.M.) and the association of state school mathematics masters (A.P.M.E.P.) have insisted on the scholastic value of interdisciplinary teaching, with activities related to the child's environment. It has been suggested to devote some of the teaching to the study of themes, and reports on experiments have been published.

All this should favour the development of statistics in secondary schools, but the introduction of new subjects and methods needs teachers with more extensive training, initially and in their careers, than is presently given to them.

The present state of affairs is likely to change, slowly, with the modification of fourth form (15 to 16 year old) syllabuses and with the promotion of educational projects. This may influence the future syllabuses in secondary schools. The educational projects, which are interdisciplinary, may be related to the child's surroundings and may include individual work or work in groups, etc., Various subjects have been tried, particularly regional economic studies, social and cultural enquiries.

Research

Since 1971 groups of teachers in France have worked or are working on the possible development of the teaching of statistics in secondary schools. This research has been organised by the national teaching research institute (I.N.R.P.) by the I.R.E.M.'s and by members of the A.P.M.E.P. Accounts of experiments have been published (see the bibliographical appendix).

Looking back on all the work done in secondary schools, since 1971, I feel that the experiments carried out can be summed up by the points made in "Mathématiques du quotidien dans les collèges, activités d'analyse des données dans le premier cycle" published in 1979 by the I.N.R.P. - I.R.E.M. research group whose progress was specially followed by P.L. Hennequin from 1972 to 1978.

This research team comprised teachers working with the I.N.R.P. in Paris, or with the I.R.E.M.'s in Bordeaux, Grenoble, Lyons, Orléans, Rennes and Rouen.

Other groups have worked more recently on statistics in secondary education, e.g. at Clermont Ferrand, but their means, methods and conclusions are not different.

The points of view adopted

In secondary schools, the intentionwas to teach "neither probability theory, nor the conditions of validity of statistical tests. On the other hand, statistics may be taught with a view to analysing data, since that can be done before pupils are acquainted with the theory of probability. Moreover, while many interesting studies have shown that (even very young) children are interested by random phenomena, in the form of games based on the raffle, it seems that this method forgets an important aspect of the statistician's activity, namely not favouring any a priori probability model, but observing data obtained in a natural context".

The data are real and numerical. At this age pupils are not afraid of, and may even "be interested by the juggling of numbers". They are beginning to be proficient in the basic operations, but are curious to use and to get the feel of large amounts of numbers, or of large numbers.

A boat of 10,000 tonnes displacement means nothing to a young pupil who has just learned this number. But after a study of the traffic in a port, or perhaps visiting a port, his reaction will be much more conscious. "Moreover, while the analysis of numerical real data is in itself a statistical activity, it can lead to a probabilistic model. This may occur through the need to predict, or through the repetition of measurements".

The reality of the data is important to these adolescents who are becoming aware of their social and cultural environments. Besides this, the ability to understand a table, to interpret a graph or to use certain properties characteristic of statistical series are necessary training for all pupils, especially those who will not be going on to further education.

Besides these goals, some teaching principles were chosen for this training: Action and communication, at the basis, can be used to develop each child's learning ability, at his own pace, or to incorporate certain statistical or mathematical ideas brought up by the questions which the groups asked themselves about the data.

These general conditions imply that there is no question of stating a theory, but that "the goal is to create situations where the following abilities are required:

- understanding information
- criticising information
- making decisions with knowledge of the chance of error
- accepting discussion and the confrontation of ideas within a group"

"This approach is also an opportunity to teach mathematics a different way. By a continual exchange between the pupil's representation of a situation and the situation itself, a place is found for mathematical structures, at the level of cognitive analysis which may be used."

Scholastic objectives may then be put in three groups:

1) To teach the pupils the general method of the statistician, i.e. unprejudiced examination of the data; their analysis with respect to a question or a problem to be solved; a critical attitude to the data, the pertinence of the question to the methods employed and to the answer obtained. This critical attitude is developed by the comparison of ideas from different groups.

2) To acquire certain habits
The pupils are expected to know

- how to draw up a table of the data
- how to ask questions about the data
- how to take the initiative to answer these questions: choice of method, calculation, graphical representation or other methods of interpretation.
- how to criticise a statistical treatment or an interpretation.

3) To develop mathematical training and activities.
"The actual mathematical side of these goals is of course essential. During the treatment of the data, pupils use their previous knowledge, may build new tools for the analysis of the data and may learn new ideas specially introduced for the occasion".

Amongst other mathematical concepts and procedures which may be envisaged are:

- classifications, mappings
- work on $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{D}$, $\mathbb{Q}$
- mean and spread

- proportionality
- reading and drawing graphs
- an approach to probabilities

Thus, without adding to the mathematical syllabuses, we have a concrete "approach to many important parts of the syllabuses".

Methodology of most of the published research

There was not any laboratory research. It was decided to work on observations of groups of children in their classes.

The I.N.R.P. group writes, "Observation is a means of obtaining an objective opinion of teaching method. Instead of just saying "That worked", or "They didn't understand a thing", by careful osbervation we may collect sufficient information to explain the childrens' reaction to a teaching method. In our view, observation is an illustration and not a proof, since the point is not to prove a general truth, but to describe a particular truth.

The goal is thus to find out what was wanted to be done, what was wanted to be observed and what was done and what actually happened".

These observations allow us to infer how the children grasped the situation.

The shortage of psychologists in the research teams forced the teachers to learn, without the supervision of psychologists, the methods of objective osbervation and rigorously to respect instructions, particularly in order to reconstruct the chronological order of the work, the expected and unexpected behaviour.

Research on 11 to 13 year olds done by the I.N.R.P. and the I.R.E.M. involved seven teams from Bordeaux, Grenoble, Lyons, Orléans, Paris, Rennes and Rouen.

Each team comprised mathematics masters and psychologists. There was freedom of choice of the subject.

Pupils worked in groups of three or four. Teachers set the work in the least restrictive way possible, such as:

"With all these data what kind of enquiries can you imagine"

or "Find a subject that interests you and write a questionnaire to be given to your classmates".

The observers, one per group, whenever possible, noted the childrens' discussions, their behaviour and the teacher's action. The team's report was a synthesis of these observations. A synthesis of the team's work was presented in the group's report.

All the teams produced a report on the subjects they studied (see biographical appendix).

Teaching situations produced in the classes

Once more, we shall use the I.N.R.P. - I.R.E.M. research as an illustration, but the results are true of what happened in other research groups.

135 ordinary urban and rural classes were involved in the I.N.R.P. - I.R.E.M. research.

"Topics studied by the teams were of general or local interest (variation of the population of a county, economic activity of a port) but were adaptable to any area, or were related to the childrens' immediate environment (surveys of hobbies, comics, television programmes, etc.,....). Topics were chosen with a view to the mathematical and statistical ideas they would involve and with a view to involving the children, e.g. by collecting data and building a table, children became more aware of the difficulty of carrying out an enquiry or a survey."

Method used when the data was supplied

Teachers distributed tables of data, generally large ones, and after a few explanations of how to read them, gave the least restrictive instructions possible for their use.

Pupils knew that the work could last several classes (sometimes a dozen hours). A period spent looking for questions about the data was followed by a class synthesis where ideas could be shared. Each group then chose a subject on which it was to work and to prepare a report.

The next stage involved work in groups. Time spent on this second part varied widely with the subject (2 hours, sometimes 4 hours). Setting out was sometimes slow, since the groups had to organise themselves, define their goals, choose secretaries, chairmen and the sharing out of the work.

Work in groups must benefit all the class and may help other groups, so at the end of this stage a class synthesis was held.

Chairmen of each group explained the aims and the work accomplished up to date.

The teacher must decide when to hold the synthesis. This is sometimes difficult to decide.

When much work is involved the above synthesis is an opportunity to review the subject, to foresee possible extensions, to decide on work to be corrected, completed or re-oriented and to reveal similarities between the work of different groups.

The third stage was often sufficient to finish the work in each group before the final synthesis which gave all the pupils the feeling of having participated in a collective activity... which they should be able to explain.

Some long subjects may need more than three sessions.

Pupils had to explain their work and to justify their actions in a way beneficial to all.

The teacher's role is above all to animate the sessions. When several or all the pupils are in difficulty, the teacher can organise information sessions, i.e. classes, specially on certain necessary mathematical tools.

Enquiries

The main difference with the above was in the first stage.

This involved several phases: choosing a subject, writing the questionnaire, preparing and carrying out the enquiry, examining the results obtained.

These phases are done or not by the pupils, depending on the teacher's intention when it was decided to work on enquiries.

Some enquiries involved a complete population of pupils, while others were based on samples and on the whole population in order to stimulate reflection on the idea of a sample.

The data

The type of data in each report is given over.

Behaviour observed

Owing to the amount of elementary work involved, to certain aims, especially the use of communication, and to the possibility of different rates of work, group work was preferred to individual work.

The advantages of group work were pointed out by all the research teams: mutual aid, spreading the work, which allows everyone to participate, the useful discipline implied by communication.

When there was argument between the leaders of a group, the teacher had a delicate part to play in solving problems of relations, or in making some groups efficient.

Pupils' interest was stimualted by the (relative) freedom in the choice of subject and the methods employed. They decided, in their own language, their topics. This freedom led them to accept responsibility, since choice implies a commitment, insisted on by the teacher, to produce a piece of work.

Report Number	Subject	Source
1	Weather (Bordeaux, Brest, Strasbourg)	Meteorological offices in Bordeaux, Brest and Strasbourg
2	Weather (Brest, Chateauroux, Orléans, Strasbourg)	Meteorological offices in Brest, Chateauroux, Orléans, Strasbourg
3	Water consumption in Grenoble	Grenoble Water Board
4	Electrical consumption in the South East from 1971 to 1975	South-East Electricity Board
5	Electrical consumption in the South-East on the third Wednesday in December*	Lyons and Marseilles Electricity Board
6	Numbers of pupils	Ministry of Education
7	Sporting performances	Files kept by the sports masters
8	Population of the Gironde and the Dordogne	l'I.N.S.E.E. (National Statistical office)
9	Football	Publications of the Fédération Française de Football
10	Car registration in the Rhône	Prefecture Lyons
11	Car registration in the E.E.C.	Motor trade paper
12	Loire river	Service hydrologique Centralisateur d'Orléans
13	Hobbies - holidays	Enquiry done by the pupils
14	Measurements of the length of a room	Done by the pupils
15	Port of Bordeaux	Bordeaux port authority
16	Port of Rouen	Rouen port authority
17	Sport results	Files kept by the sports masters
18	Surveys	Enquiries done by the pupils
19	Heights of 4th formers in a school in Orleans	enquiry done by the teachers
20	Heights and weights of 4th formers in a school in Valence	Enquiry by the pupils
21	Heights, weights and sports results	Enquiry by the pupils
22	Heights and weights from birth to 18 years	Collected by science masters from the health records
23	Television	Questionnaire sent to the pupils

* N.B: This is the last Wednesday off school before Christmas.

Number		Number of Tables	Dimensions of the largest tables	Amount of numbers given
1	Weather (Bordeaux, Brest, Strasbourg)	11	13 x 12	1,100
2	Weather (Brest, Châteauroux, Orléans, Strasbourg)	5	20 x 12	1,007
3	Water consumption in Grenoble	3	31 x 12	1,750
4	Electrical consumption in the South East from 1971 to 1975	10	31 x 6	1,795
5	Electrical consumption in the South East on the third Wednesday in December	3	24 x 6	204
6	Number of pupils	12	24 x 6	591
7	Sporting performances	9	25 x 2	250
8	Population of the Gironde and the Dordogne	4 2	15 x 9	435 270
9	Football	1	20 x 20	380
10	Car registration in the Rhône	2	8 x 12	188
11	Car registration in the E.E.C.	4	34 x 8	943
12	Loire river	10	31 x 12	3,653
13	Hobbies - holidays	0	0	0
14	Measurements of the length of a room	2	100	60
15	Port of Bordeaux	26	38 x 8	3,982
16	Port of Rouen	12	50 x 6	1,970
17	Sport results	2	34 x 5	260
18	Surveys	3	14 x 12	288
19	Height of 4th formers in a school in Orleans	1	2 x 8	16
20	Heights and weights of 4th formers in a school in Valence	8	34 x 4	800
21	Heights, weights and sports results	5	32 x 7	784
22	Heights and weights from birth to 18 years	8	39 x 4	1,070
23	Television	6	35 x 10	1,500

Most themes favoured interdisciplinary collaboration, which was indeed done in some cases. This much improved the atmosphere and everybody's interest. Questions asked by the groups required the teacher to inquire for more information from teachers of other subjects.

There were pupils, however, good scholastic ones in particular, who were not interested in their work, especially at the beginning: "But that's not mathematics!".

Lower school children, especially the 11-12 year olds, had much imagination and curiosity, so that ideas were numerous, much more so in fact than with the 16-18 year olds.

Mathematical ideas used or introduced

Sets. Cartesian product, relations, mappings, operations in P(E), partitions. Generally speaking, the set ideas were not often explicitly used by the pupils, even when they were known to them. This may be because this kind of work is often felt by the children to be cut off from any mathematical context. The opposite attitude was also seen. Some pupils wanted at all costs to find or introduce mathematical ideas.

$\mathbb{N}, \mathbb{Z}, \mathbb{D}, \mathbb{Q}$. Sums - ordering. Many questions were asked about comparisons, variations, and calculating differences.

Numerical calculations

Calculations were often repetitive, so in order to avoid boredom, pupils could use calculators, sometimes programmable ones.

When, in some teams, the repetitive numerical work had to be done again, to develop certain questions, results were obtained with a microcomputer set at the pupils' disposal (weather study in Bordeaux, when after the work on temperatures the pupils wanted to undertake a study of rainfall).

Graphical representation

"Experiments show that while most pupils have at some time of their education seen graphs, only rarely have they had to make them... Pupils often have seen and read graphs in natural science and geography classes, and in these subjects only a few classes have made graphs".

Pupils' experience in this area generally varied widely. Tests were carried out on 679 pupils to clarify this matter. The analyses of these tests point out:

"11 to 12 year olds seem to have some knowledge of the word graph. Nonetheless, this knowledge is often vague. More often than not, the work done involved labelling points...".

Even 13 year olds were rarely given the choice of representation, or even, the graduation of graphs.

The most common spontaneous graph was of the "temperature" kind, especially when time was involved.

Yet one is surprised by the number of pupils who, confronted by a problem with two variables, do not think of a graph to picture or to understand better the variations. This was also osberved among 16-18 year olds.

Difficulties in this kind of work, requiring research by the children and help from the teacher, were due to the choice of graduations (regular or not, no graduation), to the choice of origin, and to the data which sometimes were too large.

But these difficultites, the failures, the reports to the class, classmates' criticisms gradually led to more efficient graphs, richer in information. These personal contributions and successive corrections are very improving, on condition that the pupil is allowed to create these transformations himself. (See Appendix 2) "Pupils who draw a graph "to see" do not always feel the need to deduce consequences from it, or to interpret graphs which become an end in themselves". Many could profit by training in the use of graphs as a tool for analysing a question.

Mean - Mode - Median

11 to 12 year olds often know the word average without knowing what it means, in which case some pupils make its calculation an end in itself.

They often refute that an average can contain decimal places. "There can't be a car cut in two" (report; car registration).

It appears that for most pupils the mean is a number in the sequence, the most representative number, and not a real number the central characteristic of the sequence, i.e. a very abstract number in fact, produced by a formal definition.

This remains true of 12 to 15 year olds. Moreover, it seems necessary to familiarise the pupils with the average.

Mode and mean are sometimes mixed up. The median is quite frequently used. "What if we took the middle pupil" (report; sports results).

The mode and above all the median are much more concrete measures than the mean.

The mean certainly is difficult for children, e.g. the mean of the means, cases where the same number appears several times, or where null values are forgotten, or mixing up comparison of the means and comparison of the size of the populations.

Ratio - Frequency - Proportionality

These ideas occur spontaneously in statistical calculations whenever comparisons are required.

Yet, "reasoning on ratios and a fortiori the use of percentages is far from easy to 11-15 year-olds... as no doubt to many grown ups. This kind of reasoning is considered by Piaget and Inhelder as one of the characteristics of formal reasoning which begins at 11 to 12 years". Complete mastery of these ideas is achieved only slowly and at different personal rates.

"All the evidence shows that frequency and percentage are difficult ideas. They are assimilated only slowly, through contact with many rich situations where the pupil can ask himself questions. Current exercises, which put too much accent on the calculation, do not favour mastery of the idea". It appears that for 11 to 13 year olds calculation of frequencies as approximate decimal fractions is sufficient.

Spread

Certain themes required groups to compare variations and to deduce their consequences, e.g. the work on weather.

The reports say little on the small interest given by the children to spread and in particular to its representation by a number.

The average of a set of numbers seems natural enough, but bothering with the average deviation from the average and with a number to represent it require model building and abstraction beyond most 11 to 14 year olds. The idea of distance, moreover, is not properly mastered.

Some themes led to drawing scatter diagrams (report on heights, weights and sporting ability). Groups took account of the cloud of points associated with a sequence of two variables by drawing a line with as many points above it as below. (This idea was often suggested by the teacher). Pupils then drew two lines parallel to the first, enclosing the cloud of points and even rejecting some far flung ones. Nearly all the points lay in this band (92%, report on sporting ability).

The talk given to the class by this group had a large impact, a sign that pupils are favourable to some elementary modelling of real phenomena.

Correlation

Study of statistics in secondary schools can lead to consciousness of correlations between variables. Three of the themes used by the I.N.R.P. - I.R.E.M. research gave rise to studies çorrelation:

In the study "Heights-weights-armspread" pupils spontaneously sought out

"Are the tallest the heaviest?"
"Do height and armspread favour putting the shot?"

In the study of sporting ability, we found the question, put by the pupils,

"Is running related to the long jump? to the high jump?"

Finally, in the study of the port of Bordeaux, the variation with time of the number of boats coming in and their laden weight led pupils to study the relation between the two and to seek why the first was decreasing and the second increasing.

The methods used were graphical (comparison of several graphs, study of scatter diagrams) or tables of spreads.

Awakening of awareness to random phenomena and prediction

In these experiments, the childrens' interest was always very strong for any prediction which they were told they would be able to check.

"But extending in time the results obtained from a descriptive study was not always spontaneously suggested". At this age, we can try only to make the children sensitive to the problem, in view of a later introduction of probabilities.

Some mathematical tools (absolute value, average, frequency, spread, elementary combinatorial analysis,...) are not sufficiently mastered, besides which setting up even an elementary model requires formalisation which the children often find difficult.

They are sensitive to the gaming aspect of prediction, and ready to propose ideas. What was done was considered to be useful. Pupils, aided by the teacher, became aware that very often measurement of some quantity is associated with a random variable, that any future experiment involves risk, but that randomness generally ensures that the results lie in some interval [m, M] determined by analysis of the results already obtained.

Some examples are:

1) Sampling

Sampling was approached through prediction. One pupil proposed: "If I question only 6 out of 35 pupils, will I have everyone's opinion?".

Study of a sample and then of the entire population was strongly motivating.

2) Treatment of the data

"How many French made cars will be registered in 1975?"
(report on car registrations)

"Between what two numbers does the length of the class lie?"
(report on the measurement of a length)

"Can a pupil's speed be used to predict his long jump?"
(study of sporting ability)

Methods used called for the mean, the extent of a sequence and graphical methods (lines illustrating tendencies, scatter diagrams).

During the discussions in the syntheses, many pupils were aware that prediction is the use of a model (a very simple one of course at this stage), a formalisation using elements of the statistical study: mean, percentage, tendency lines, the interval [m, M].

CONCLUSION

We have discussed here research done by the I.N.R.P.-I.R.E.M. from 1973 to 1978, but other more recent work exists also, e.g. Clermont-Ferrand 1980-81, Rennes 1978-79, suggestions for adult teaching adults in Lilles, but which could be adapted to secondary school use, research on graphs (interdisciplinary, not just statistical) done by an I.N.R.P.-I.R.E.M. group.

All the conclusions corroborate what we have said above, particularly:

Secondary school children are not at all put off by large amounts of information. Indeed, it is a situation resembling closely their everyday experience.

They are strongly motivated by real data, which they may have collected themselves, or which deals with their social, economic or cultural environment.

They are ready to undertake long projects with many calculations, provided they are clear of the aims and if they have some equipment at their disposal as well: calculators or microcomputers.

They are able to and interested in working on random phenomena, in foreseeing possible errors, in predicting. It would thus be a shame to restrict them to purely deterministic teaching whereas they live in a world where statistics and probability play a large part. They will often have to make decisions. Our duty is to give them the tools to make these decisions conscious, calculated and involving a minimum of risk.

References

[1] Mathématiques du quotidien dans les collèges (Activités d'analyse des donnés dans le premier cycle). ed.: l'Institut Notional de recherche Pédagogique.

[2] Dossiers de la recherche I.N.R.P.-I.R.E.M.

Nos.	TITRE DU DOSSIER	EDITE PAR
1	Climatologie (Bordeaux, Brest,Strasbourg)	I.R.E.M. de Bordeaux
2	Climatologie (Brest, Châteauroux, Orléans, Strasbourg)	I.R.E.M. d'Orléans
3	Consommation d'eau à Grenoble	I.R.E.M. de Grenoble
4	Consommation d'électricité dans le Sud-Est de 1971 à 1975	I.R.E.M. de Lyon
5	Consommation d'électricité le 3e mercredi du mois de décembre dans le Sud-Est	I.R.E.M. de Lyon
6	Effectifs scolaires	I.N.R.P. Paris
7	Étude de performances sportives	I.R.E.M. de Bordeaux
8	Évolution de la population en Gironde et en Dordogne	I.R.E.M. de Bordeaux
9	Football	I.N.R.P. Paris
10	Immatriculation de voitures dans le département du Rhône	I.R.E.M. de Lyon
11	Immatriculation de voitures dans les pays de la C.E.E.	I.R.E.M. de Rennes
12	Loire	I.R.E.M. d'Orléans
13	Loisirs - Vacances	I.R.E.M. d'Orléans
14	Mesures de la longueur d'une salle	I.R.E.M. de Bordeaux
15	Port de Bordeaux	I.R.E.M. de Bordeaux
16	Port de Rouen	I.R.E.M. de Rouen
17	Résultats sportifs	I.R.E.M. de Grenoble
18	Sondages	I.R.E.M.Paris-Nord
19	Taille d'élèves de 5e d'un C.E.S. d'Orléans	I.R.E.M. d'Orléans
20	Taille et poids d'eleves de 5e d'un C.E.S. de Valence	I.R.E.M. de Grenoble
21	Taille, poids, envergure et résultats sportifs	I.R.E.M. de Rennes
22	Tailels et poids: croissance de la naissance à dix-huit ans	I.R.E.M. d'Orléans
23	Télévision	I.R.E.M. de Rouen
-	Autres dossiers	

[3] Etude d'un thème statistique en classe de 5ème-6ème : étude de la vente de journaux à St. Brieux (I.R.E.M. de Rennes)

[4] Dossiers statistiques (cotes et débits du barrage de Rochebut (Allier)) I.R.E.M. de Clermont

[5] Extraits du travail de recherche spontanée en 6ème. I.R.E.M. de Clermont

[6] Activités mathématiques. Centre Université Economie d'éducation permanente (C.U.F.E.P.) Lille

[7] Matemactives. I.R.E.M. de Lille

[8] Publications de deux groupes de stagiaires (en préparation). I.R.E.M. de Montpellier

BEST COPY AVAILABLE

TEACHING STATISTICS CONCEPTS THROUGH INSTRUCTIONAL GAMES

**George W Bright, Northern Illinois University, John G Harvey, University of Wisconsin and Margariete M Wheeler, Northern Illinois University*

Since 1976 the authors have been engaged in a systematic investigation of the cognitive effects of mathematics instructional games. From the outset the research has focused on identifying effective uses of games in order that recommendations for appropriate incorporation of games into mathematics instruction can be made.

A game is defined by seven criteria:

1. A game is freely engaged in.
2. A game is a challenge against a task or an opponent.
3. A game is governed by a definite set of rules. The rules describe all procedures for playing the game, including the goal(s) sought; in particular, the rules are structured so that once a player's turn comes to an end, the play just completed cannot be retracted or exchanged for another move.
4. Psychologically, a game is an arbitrary situation clearly delimited in time and space from real-life activity.
5. Socially, the events of the game situation are considered in and of themselves to be of minimal importance.
6. A game has a finite state-space (Nilsson, 1971). The exact states reached during play of the game are not known prior to the beginning of play.
7. A game ends after a finite number of moves within the state-space.

The first five of these criteria, based upon a discussion of play, games, and gaming situations (Inbar and Stoll, 1971), give formal, psychological, and sociological characteristics of games. The last two criteria were evolved in order to exclude other, related activities or structures such as play and puzzles.

An instructional game is a game for which a set of instructional objectives has been determined; these instructional objectives may be cognitive or affective and are determined by the teacher(s) planning instruction. The source of ideas for the statistical instructional games used in the research was Developing Mathematical Processes (DMP) (Romberg, Harvey, Moser and Montgomery, 1974, 1975, 1976). DMP is a rich source of mathematics instructional games for elementary school students. These games frequently also seem easily adaptable for other ages and for more sophisticated content.

The notion of cognitive effects derives from the work of Bloom (1956). In describing a taxonomy of educational objectives, Bloom (1956) states that the cognitive domain "includes those objectives which deal with the recall or recognition of knowledge and the development of intellectual abilities and skills" (p.7). Thus, cognitive effects are effects which describe changes or differences in the attainment of objectives in the cognitive domain. For the studies reported here, the cognitive effects are defined as the change in attainment of an objective in the cognitive domain due to a particular instructional treatment; that is, due to the playing of mathematics instructional games.

During 1979-82 a series of studies has been conducted to examine the effectiveness of mathematics instructional games in each of several combinations of instructional level and cognitive level. Three instructional levels have been identified: post-instructional level, co-instructional level, and pre-instructional level. An instructional activity is at the post-instructional level if it is used after instruction designed to provide students with mastery of the content; an activity is at the co-instructional level if it is used in conjunction with instruction designed to produce mastery; and, an activity is at the pre-instructional level if it is used prior to such instruction. Hence, instructional level is dependent on the prior instruction provided to a particular group of students.

Cognitive level is defined in terms of Bloom's 1956 taxonomy of six levels: knowledge, comprehension, application, analysis, synthesis, and evaluation. Of these six levels, evaluation essentially never appears in U.S.A. school mathematics curricula; synthesis appears only very rarely. Because the research conducted was limited by school curricula in the U.S.A. and was not designed to change those curricula, the evaluation and synthesis cognitive levels were omitted from consideration.

The remaining four cognitive levels and the three instructional levels yield twelve combinations; within each of these combinations it is possible to study the effectiveness of games as teaching tools. However, one of these combinations, namely pre-instructional/analysis, does not seem appropriate for the use of games. In this combination, students would be expected to use content at the analysis level prior to instruction designed to produce mastery of the content. While some students might succeed at this activity because of well-developed problem solving skills, most students should not be expected to use content at this high cognitive level prior to instruction. Hence, this combination was also omitted from further consideration.

For each of the remaining eleven combinations, studies at two different grade levels were conducted during 1979-82 with a variety of mathematics content. Of these twenty-two studies, four dealt with probability and statistics concepts with students aged 12 to 15 years. The distribution of these four studies among the eleven combinations was as follows:

1. pre-instructional/knowledge: one study
2. co-instructional/comprehension: one study
3. pre-instructional/application: two studies

Three of these studies are reported in this paper; the fourth, one of the pre-instructional/application studies, is in progress.

Several assumptions underlie these studies:

1. The mathematics content was substantial and would not be learned without instruction specifically designed to teach that content. Hence, control groups were not consistently used in these studies. Specifically, in the pre-instructional studies it was not necessary to have a control group in order to measure the effectiveness of a game. In the co-instructional study it was of course vital to have such a control group, since regular instruction was also occurring.
2. At both the pre-instructional and co-instructional levels, prior to the start of the study the students had not received instruction designed to produce mastery of the content.
3. Games could be found or modified which would provide instruction of the mathematics content at the appropriate cognitive level.

GENERAL RESEARCH DESIGN

The same basic design was used in all of the studies. Each study was based on the instructional day; that is, any school day during which instruction on the particular content of the game was given and/or the game was used. Games were played for eight instructional days. In pre-instructional studies these eight days were spread across four consecutive weeks; that is, the games were played twice per week for 20 minutes on each day of play. For the co-instructional study the eight days augmented standard instruction, so the game playing days were unevenly spaced over a three-week period. Again, on each of the eight days, the games were played for 20 mintues per day. In each study, during the week prior to the game playing pretests were given, and during the week after the end of game playing, posttests were given. Cognitive effects were measured by comparison of pretest and posttest scores in the pre-instructional studies and by analysis of covariance on the posttest scores with pretest scores as the covariate in the co-instructional study. As noted earlier, in the pre-instructional studies control groups were not used. In the co-instructional study, however, a sample of six students from each classroom was randomly selected, subject to a sex distribution approximately equal to that of the whole class, to serve as a control group. Control subjects received the regular instruction provided by the teacher, but they played

Mastermind (Invicta, 1972) instead of the particular statistics game being studied.

Also, in each study the four-card test of formal operations (Adi, Karplus and Lawson, 1980) is given as a pretest. Students aged 12 to 15 may be in transition between concrete and formal operations, and the particular level of skill that they have in dealing with material requiring logical reasoning may affect their success with that material. Hence, it was decided to measure formal operations so that content scores could be adjusted for formal operations skill.

PRE-INSTRUCTIONAL STUDIES

Two studies were conducted at the pre-instructional level. One dealt with knowledge of fair/unfair, and one dealt with the application of probabilities in decision making.

For the fair/unfair study eight pairs of games (Bright, Harvey and Wheeler, 1981) were used. These games were adapted from Developing Mathematical Processes. The content of these games is at the knowledge cognitive level. Subjects were 36 students in grade eight (about age 13).

A 28-item content test, consisting of two 14-item subscales, was given as both a pretest and a posttest. In each of the 28 items two games were described and students were asked to determine the fairer of the pair of games. For each item on one subscale students were given data summarizing 50 plays of each of the two games. (These data were generated by computer simulations of the games). The items on the other subscale were not accompanied by these simulated plays. A sample item is provided in Bright, Harvey and Wheeler (1982).

The means and standard deviations and the comparisons of the pretest and posttest scores are presented in Table 1. Analysis of variance of the subscale without simulation data given, with sex as the classification variable, yielded a significant sex-related effect ($F = 6.62$, $p < .015$) favoring the males. Analyses of covariance, with the four-card test and the pretests as covariates, yielded no significant effects.

TABLE 1

Pretest - Posttest Comparisons for Fair/Unfair Games at Grade 8

Scale	Pretest Mean (SD)	Posttest Mean (SD)	t	p
14 items with data given	6.50(2.27)	7.56(2.30)	3.83	.001
14 items without data given	6.44(1.78)	5.81(2.12)	-2.05	.048
Total test	12.94(3.22)	13.37(3.91)	0.93	.357

The games seemed to cause an improvement in performance at the knowledge cognitive level when data summarizing plays of the games were available. The marginal negative effect on the scores of the subscale without data is inconsistent with the results of a similar study at grade six (Bright, Harvey and Wheeler, 1982) and in the larger context of this project seems to be a spurious result. It is encouraging that in this setting students seemed able to learn to interpret data through the playing of the fair/unfair games.

It should be noted that in an earlier study (Bright, Harvey and Wheeler, 1980b) the same games were used with 82 grade seven students (about age 12). Although not identified at the time, the instructional level was pre-instructional. Two types of game-playing groups were formed: heterogeneous groups and homogeneous groups. Based on pretest scores students were grouped into four strata of equal size. Heterogeneous groups of size four were formed by randomly choosing one student from each stratum. Homogeneous groups of size three or four were formed by randomly choosing students from the same or adjacent strata. Thirteen heterogeneous and twelve homogeneous groups were created. Students who did not complete the posttests were dropped from the analyses. The pretest and the posttest were like the subscale reported earlier without any simulation data presented with the items.

The analyses of the data are presented in Table 2. Analysis of covariance with the pretest as the covariate was conducted on the posttest scores with the treatment and sex as classification variables. None of the F-values were significant at the .05 level; that is, neither treatment nor sex differentiated performance level.

TABLE 2

Pretest - Posttest Comparisons for Fair/Unfair Games at Grade 7

Grouping	N	Pretest Mean (SD)	Posttest Mean (SD)	t	p
Heterogeneous	49	16.0(5.1)	17.7(4.6)	3.01	.004
Homogeneous	33	15.7(4.9)	16.9(4.6)	2.82	.008

The second pre-instructional study dealt with application of probabilities in making decisions. Number Golf, the game used in this study, was based on a game described by Bright (1980). In this game a goal number is designated and players try to reach this goal number by adding or subtracting the sum of two numbers generated by dice rolls to their cumulative scores. The content of Number Golf is at the application cognitive level. Subjects were 74 students in grade seven (about age 12).

For this study three tests were used. One test was given only as a pretest; it was a 10-item test in which students were told that a pair of regular dice had been thrown 36 times; students were asked to predict how many times each of 10 given sums would appear. Students were given two points for each answer that matched exactly the predicted number of occurrences of each sum and one point for each answer that differed from the predicted number by only one. Two tests were given both as pretests and as posttests. The first was a 20-item test; in each item students were asked to choose the most likely of three sums generated by the toss of a pair of dice. Ten of these items dealt with standard dice; ten, with non-standard dice. The second was a 42-item test; for each item a goal number, a given score, and a number designated as the sum of a throw of two dice were given. Students were asked to choose the better of two possible moves, generated by adding or subtracting the given sum to the given score, in order to reach the goal number in as few moves as possible. Students could also indicate that the two moves were equally good. A sample item is given in Figure 1. Twenty-one of the items dealt with standard dice; 21, with non-standard dice.

FIGURE 1

Sample Test Item for *Number Golf*

Goal Number:	12
Given Score:	9
Dice Roll:	5

Which move would be better?

a. Subtract 5 from the Given Score to get a total of 4.
b. Add 5 to the Given Score to get a total of 14.
c. It doesn't matter which move is made.

The means, standard deviations, and comparisons of pretest and posttest scores are presented in Table 3. Analysis of covariance on the most likely sum/regular dice test with the four-card test, the number of occurrences test, and the pretest of most likely sum/regular dice as covariates and with sex as the classification variable yielded a significant sex effect ($F = 6.85$, $p < .011$) favoring the males. Analyses of covariance on the other three posttests with analogous sets of covariates and sex as the classification variable yielded no significant effects.

TABLE 3

Pretest - Posttest Comparisons for Number Golf Games at Grade 7

Scale	Pretest Mean (SD)	Posttest Mean (SD)	t	p
Number of occurrences[a]	6.94(3.95)	not given		
Most likely sum[b]				
regular dice	2.61(2.30)	4.42(2.12)	6.80	.000
non-standard dice	1.99(1.70)	3.38(2.09)	4.67	.000
Choosing move[c]				
regular dice	6.92(2.29)	7.27(2.94)	1.19	.240
non-standard dice	6.10(2.10)	6.00(2.50)	-0.31	.754

a. 10 items

b. 10 items in each subscale

c. 21 items in each subscale

The most likely sum test is not at the application cognitive level but is rather at the comprehension cognitive level for students who have not yet received instruction on probability. Hence, Number Golf was not effective at producing improvement in application cognitive level performance among the seventh graders who participated, though performance at a lower cognitive level did improve. A replication of the study with grade nine (about age 14) students has just been completed. Data should be analysed by mid-summer 1982.

CO-INSTRUCTIONAL STUDY

One co-instructional study dealt with the concept, average. This study was at the comprehension cognitive level. The game, Average Hands, was taken from Developing Mathematical Processes. In the game students are dealt 12 cards from a deck of number cards. Players must arrange their cards into three subsets: one with three cards, one with four cards, and one with five cards. After all of the arrangements are completed, players compare the averages for each size of subset. The player with the highest average scores 3, 4 or 5 points, depending on the size of the subsets being compared. Variations of this game were developed by changing the scoring rule to reward the smallest average in a comparison and by varying the numbers occurring in the number decks. Subjects were 102 grade seven students (about age 12) with 80 students in the experimental (game playing) group and 22 in the control group.

A variety of tests were used. Five speed tests were given only as pretests: (a) a test of all 100 multiplications basic facts, (b) a test of 12 addition items, with four to eight two or three-digit addends in each item, (c) a test of 18 long division items, with one and two-digit divisors, (d) a test of 8 rounding items, with each item being a mixed number to be rounded to the nearest whole number, and (e) a test of 7 multiple-choice estimation of number items, each of which was a "word" problem with five approximate answers given. Two tests were given as pretests and posttests. One was a test of 20 averaging items, each of which was a set of four to eight one-, two-, or three-digit numbers for which an average was to be computed. The other test contained 20 items; for each item a set of nine numbers, in the range 1 to 99, was given, and students were asked to choose two disjoint subsets of three or more numbers each, so that the averages of the subsets were within specified ranges. A sample item is provided in Figure 2. The sets of nine numbers were randomly generated and the ranges of the averages sought were chosen so that each item had one or more solutions.

FIGURE 2

Sample Test Item for Average Hands

In each of the questions which follows you are given a list of nine numbers. Using the numbers in each list, you should make two sets of three or more numbers so that the average of the numbers in Set One is between the numbers given for that set, and the average of the numbers in Set Two is between the numbers given for that set. Do not use any number in the list more than once. You do not have to use all the numbers in the list. Usually there will be several correct ways to choose the two sets of numbers.

Numbers: 1,2,3,4,5,6,7,8,9
Set One: ________________ average between 2 and 4
Set Two: ________________ average between 6 and 8

The means and standard deviations of the tests are given in Table 4. Analysis of co-variance on the posttest of averaging, with the four-card test, the addition test, the division test, and the pretest of averaging as co-variates and with the treatment group as the classification variable yielded a treatment effect (F = 5.1, p < .030) favoring the experimental group. Analysis of covariance on the posttest of averaging with the same four co-variates and with both treatment group and sex as classification variables yielded a treatment effect (F = 5.3, p < .025) but did not yield either a significant sex effect or a significant treatment by sex interaction. Analysis of covariance on the posttest of choosing sets, with the four-card test, the rounding test, the estimation test, and the pretest of choosing sets as co-variates yielded no significant effects for treatment, sex, or the interaction.

TABLE 4

Means (Standard Deviations) for Average Hands Games at Grade 7

Scale	Experimental Group		Control Group	
	Pretest Mean (SD)	Posttest Mean (SD)	Pretest Mean (SD)	Posttest Mean (SD)
multiplication	99.10 (3.1)		99.62 (0.8)	
addition	10.60 (1.7)		10.71 (1.3)	
division	16.58 (1.9)		16.71 (1.3)	
rounding	6.21 (2.2)		5.62 (2.6)	
estimation	3.95 (1.4)		3.24 (1.6)	
averaging	7.03 (3.6)	9.75 (3.7)	6.36 (3.1)	8.27 (3.0)
choosing sets	6.58 (4.5)	12.46 (5.7)	5.46 (3.9)	11.36 (6.2)

The game seemed to enhance regular instruction on computing averages. This behavior, however, is marginally at the comprehension cognitive level. A reasonable argument could be made that for age 12 students this behavior is only at the knowledge level, since the algorithm for computing averages involves subskills that are presumably well-known to these students. Separating a set of numbers into subsets with averages within specified bounds is certainly at the comprehension cognitive level, but the game did not seem to cause any improvement in the performance of that skill beyond that provided by the regular instruction. It is also quite likely that the teacher did not expect mastery of that skill.

CONCLUSIONS

The three studies have the common thread that the students involved had not previously been exposed to instruction designed to produce mastery of the particular statistical concept or skill needed to play the particular game. All of the games seemed effective at improving performance, either alone or in addition to the regular classroom instruction, at the knowledge cognitive level. The games did not seem effective at the comprehension or application cognitive levels. In part, this may be explained by the fact that students started the experimental period with essentially no knowledge, or at least very little formal knowledge, about the content. Hence, students may have had to build all of the cognitive structures required to absorb the content within the relative brief treatment period of approximately three hours spread across eight instructional days. This may be an unreasonable expectation.

The implication of this observation for teaching is that one game, or one variety of games, probably should not be expected to carry the brunt of initial teaching of concepts like those encountered in probability and statistics. Mathematics instructional games typically focus on a limited range of concepts or skills within a limited range of cognitive levels. Initial teaching of a mathematics concept may require broad coverage of many aspects of that concept. One type of game should not be expected to provide such broad coverage. Also, for probability and statistics in particular, basic concepts may be contaminated by students' inaccurate perceptions of real-world experiences. It seems important to help students not only identify their inaccurate perceptions but also give up their confidence in those perceptions. Perhaps highlighting the important attributes of concepts prior to using games, would effectively assist understanding of probability and statistics concepts.

Previous research (Bright, Harvey and Wheeler, 1979, 1980a) had indicated that after a concept has been taught to mastery, games are effective materials for the practice and reinforcement of performance. Performance at comprehension and application cognitive levels may depend heavily on well-developed skills at the knowledge cognitive level. For games used at the pre-instructional level this background of knowledge cognitive level performance may not be strong enough to support the higher level work. It may be unrealistic to expect an eight-day treatment to provide enough time to build both the lower level skills and the higher level skills.

Future research should examine the kinds of skills and concepts that particular students develop while playing a game at the pre-instructional level as well as the different effects that may accrue with more rounds of play of such a game. Dennis, Muiznieks and Stewart (1979) indicated in their examination of the games literature that no studies examined the effects of varying the number of rounds of play of a game. Especially in pre-instructional settings where development of skill would be expected to be slow, such studies could prove of great value in the development of guidelines for the use of games.

References

[1] Adi, H., Karplus, R. and Lawson, A.E. (1980) Conditional logic abilities on the four-card problem: Assessment of behavioral and reasoning performances. Educational Studies in Mathematics, 11, 479-496.

[2] Bloom, B.S. (Ed) (1956) Taxonomy of educational objectives. New York: McKay.

[3] Bright, G.W. (1980) Game moves as they relate to strategy and knowledge. Journal of Experimental Education, 48, 204-209.

[4] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1979) Using games to retrain skills with basic multiplication facts. Journal for Research in Mathematics Education, 10, 103-110.

[5] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1980a) Using games to maintain multiplication basic facts. Journal for Research in Mathematics Education, 11, 379-385

[6] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1980b) Achievement grouping with mathematics concepts and skills games. Journal of Educational Research, 73, 265-269.

[7] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1981) Fair games, unfair games. In: A. Schulte and J. Smart (Eds) Teaching statistics and probability, 1981 yearbook of National Council of Teachers of Mathematics. Reston, VA: NCTM, 49-59.

[8] Bright, G.W., Harvey, J.G. and Wheeler, M.M. (1982) Using games to teach some probability concepts. Paper presented at the First International Conference on the Teaching of Statistics, Sheffield, England.

[9] Dennis, J.R., Muisnieks, V.J. and Stewart, J.T. (1979) Instructional games and the computer-using teacher. Urbana, IL: University of Illinois (ERIC Document Reproduction Service No. ED 183 189).

[10] Inbar, M. and Stoll, C.S. (1970) Games and learning. Interchange, 1(2), 53-61.

[11] Invicta Plastics Ltd. (1972) Mastermind - Regular, Oadby, Leicester, England: The author.

[12] Nilsson, N. (1971) Problem solving methods in artificial intelligence. New York: McGraw-Hill.

[13] Romberg, T.A., Harvey, J.G., Moser, J.M. and Montgomery, M.E. (1974, 1975, 1976) Developing mathematical processes. Chicago: Rand McNally.

A PRACTICAL APPROACH TO STATISTICS

R Kapadia, Polytechnic of the South Bank

INTRODUCTION

Open any newspaper or magazine and one is likely to find some statistical information. Probabilistic words like chance, odds or likely are part of everyday language. Data is collected pertaining to many and various aspects of life. Decisions made by governments and industries rely heavily on published figures. Many subject areas have become increasingly quantitative. It is clear that pupils are growing up into a world becoming more dominated by statistical ideas and it is up to education to take up the challenge by providing suitable courses for all children. This paper considers the needs of ordinary children who will complete their education at age 16 or below. An attainable scheme of work, which takes into account the statistical awareness required by adults both in society and in their employment, is presented.

The first part provides a perspective showing how the influence of the subject has grown. From this the main aims are advocated and a specimen syllabus linked to applications is outlined. Suitable teaching strategies are described, illustrated by a specific example. References to suitable teaching materials are also given. Throughout, the international context is borne in mind, for statistics is an essential part in any education system.

HISTORICAL PERSPECTIVE

In a sense statistics began when man started counting and keeping records. This is a rather naive view in that, though records have been kept for many centuries, they were rarely interpreted, analysed, or more critically, used to make inferences. This had to await the emergence of probability, which was a slow process.

For man held a deterministic view of the world over a long period of time. The search for pattern and order was guided by a casual point of view. Science used empirical methods which were firmly based on cause and effect. If an experiment was repeated in the same conditions, it was assumed that the same results would ensue. This deterministic viewpoint is now being superseded especially at the microscopic level, by theories based on probability.

Probability, as a mathematical discipline is thought to have originated from a (characteristically) chance communication about gambling by Chevalier de Mere to Pascal. The study of inferential statistics began later through the work of Graunt and Petty in the seventeenth century on medical records and life insurance.

In teaching probability and statistics, we must not forget the long time the subjects took to develop and the radical change of philosophy they entail. For, according to Kendall (1970) "the late emergence of probability calculus was due to some more fundamental factor. The very notion of chance itself, the idea of natural law, the possibility that a proposition may be true and false, in fixed relative proportions, all such concepts are nowadays so much part of our common routines of thought that perhaps we forget that they were not so to our ancestors. It is in basic attitudes towards the phenomenal world, in religious and moral teaching and barriers, that I incline to seek for an explanation of the delay".

In a perceptive analysis presented at a colloquium on the history of statistics Hacking (1981) asserts that, though it seems only to provide information, statistics is part of the technology of power in a modern state. For social facts have simply become facts which are statistical in character so that 'many of the modern categories by which we think of people and their activities were put in their place by an attempt to collect numerical data'.

It was during the nineteenth century it became accepted that 'real chance exists and is part of the underlying structure of the world'. Hacking cites the following events which led to this:

(a) the theory of probability was developed,
(b) the avalanche of numbers revealed an astonishing regularity in the statistics of crime, suicide, epidemics and mathematicians attempted to analyse such phenomena,
(c) the normal curve was postulated as a universal model for human, social and biological traits,
(d) social reformers thought that one could recognise the 'boundary conditions' under which a population was governed by statistical laws,
(e) statistical mechanics and thermodynamics heralded the intrusion of chance into physics,
(f) Darwin and Galton introduced probabilistic notions into biology.

Hacking goes on to assert that 'almost no domain of human enquiry is left untouched by the events that I call the avalanche of numbers, the erosion of determinism and the taming of chance'. Paradoxically, in a quantitative universe, exactness has become impossible.

Thus, it is clear that statistics has become a vitally important subject. It affects the way we think about ourselves by determining classifications. Even birth and death, the most individual of events, are governed and described by probabilistic laws. Thus it is vital (in the literal sense too!) that children be taught statistics in school.

MODERN APPLICATIONS

One could write many pages in describing the many and varied recent applications of statistics; to give some idea of the scope just a few examples are described, according to the classification in Tanur (1972). Four major subject areas, where statistical reasoning is being increasingly applied, are considered; these are medicine, science, politics and everyday life.

Statistics has played a vital part in the progress of medical science. Substantial tests preceded the introduction of the Salk poliomyelitis vaccine, which has led to the virtual eradication of this terrible disease. A variety of drugs have been produced to combat a wide range of illnesses. Their development has required careful statistical screening. Probability has played a crucial role in many sciences. Quantum physics is firmly based on probabilistic notions, as enshrined in the famous Heisenberg Uncertainty Principle. Even life is now thought to have been caused by a chance mutation.

One finds evidence of statistics in politics constantly. Opinion polls are the popularly known manifestation yet published figures and trend analysis are used to help in making more informed decisions on a wide variety of issues.

This overlaps with the many applications of statistics in the social world - communication, leisure, national data to name just a few areas. Thus statistics affects each citizen. All major countries conduct a national census to build up a basic picture; it is important that pupils understand the rationale behind this massive exercise in data collection and the safeguards enforced to ensure confidentiality. Prices are continually rising yet people do not understand how the retail price index works. [A recent survey, published in the Cockcroft Report (1982) showed that over 50% of adults do not understand the rate of inflation]. Bar charts, pie diagrams and pictograms are often used in literature; there are also striking examples of abuses which children need to be aware of. Manipulation of figures can make a substantial difference in pay negotiations.

It is clear that statistics plays a vital role in everyday life for everyone, both at home and at work. This is the fundamental reason why the subject needs to be taught to all pupils.

SCHEME OF WORK

Having made the case for teaching statistics to all pupils one must consider the aims of a course to meet the general needs. Two aims are proposed, following the work of the Schools Council Project on Statistical Education. The author was one of the team who undertook the research for this project. The aims represent the two aspects of statistics, which is defined to be the science of numerical data. Pupils need to become aware of range of applications to which statistical methods can be applied.

(1) Children should become aware of, and appreciate the role of statistics in society. That is, they should know about the many and varied fields in which statistical ideas are used, including the place of statistical thinking in other academic subjects.

(2) Children should become aware of, and appreciate, the scope of statistics. That is, they should know the sort of questions that an intelligent use of statistics can answer, and understand the power and limitations of statistical thought.

This attitude to statistics teaching implies that pupils should deal with statistical ideas in a variety of practical situations in order to experience the ubiquity of the subject. These examples need to be chosen from different parts of the school curriculum and also from important areas of everyday life or future work. Thus the two aims are pursued concurrently so that children are introduced to the techniques and methods of statistics via realistic examples. Clearly this is a very different philosophy to the traditional, rather arid approach where statistical techniques are divorced from real data. The modern approach of teaching within context fits much more naturally with the movement towards the introduction of problem-solving in the classroom.

The statistical techniques to be studied can be divided into five categories: data collection; data tabulation and representation; data reduction; probability; inference (note the emphasis on data). These categories need to be studied in the following broad areas of application: mathematics; humanities; sciences; social sciences; other. The five areas have been deliberately chosen to fit in, as far as possible with the normal school curriculum. Nevertheless the last area is by no means the least important: it includes important applications of statistics such as medicine, management, quality control and everyday life which do not fit the neat compartments of the school curriculum.

Thus teaching materials need to be devised according to a two way classification, whereby units of work develop concepts within specific areas.

For example one could choose an application from the sciences, say the plight of blue whales. This can be used to introduce the concept of sampling and show some of the problems inherent in obtaining data. Appropriate measures of location can be applied, perhaps even a probability distribution. One can then interpret the results and make relevant inferences.

This example shows how the approach described above can be achieved in practice. It starts with a problem (How many whales are there?) and develops the necessary statistical techniques (sampling, simple inference) to help to solve it (as far as possible).

Clearly care needs to be taken in the choice of the problem and the development of statistical techniques. One needs to consider the background of the children both in terms of their cultural outlook and statistical knowledge. Some examples may be appropriate in certain countries but not in others. A selection of materials adopting this approach needs to be developed. A start has been made [7].

To ensure the breadth and scope of statistics is properly covered one needs a syllabus and list of core applications. This is presented below but the above proviso needs to be borne in mind. In some countries one would need to make appropriate adaptations. Also what is presented below is intended to be taught to all pupils, particularly those who will complete their education at age sixteen or earlier. The needs of those who will continue their study later will be much wider [8 p.71].

Syllabus

This covers the broad areas: the depth of treatment would vary according to the pupils and choice of applications.

(a) Uses of a census
(b) Types of data
(c) Sampling
(d) Obtaining data
(e) Tabulation of data
(f) Pictorial representation of data
(g) Measures of location
(h) Measures of dispersion
(i) Index numbers
(j) Correlation
(k) Intuitive understanding of probability
(l) Reading and interpreting tables or diagrams
(m) Comparing data
(n) Detecting trends

A more detailed list of objectives may be found elsewhere [8, p.68].

Core Applications

Below is a list of the important applications of statistics which all children should meet.

- (a) Census of Population
- (b) Retail Price Index
- (c) Opinion polls
- (d) Health - smoking, vaccination
- (e) Population growth
- (f) Measures used in society e.g. housing, crime, leisure
- (g) Conservation and pollution
- (h) Advertising
- (i) Errors and accuracy
- (j) Weather prediction

It is important to remember that the syllabus and applications should be considered jointly. A technique is important in statistics in terms of its applications. Thus each of the areas listed in the syllabus will be met within one of the applications; indeed some areas will occur in several applications. This will clearly be beneficial for pupils will see statistics as an interesting, live subject relevant in a variety of disciplines. Clearly one has to be careful not to introduce too many new ideas simultaneously; reinforcement is also necessary. This interlinking of techniques and applications will make statistics an exciting and meaningful area of study with direct relevance to both everyday life and future work.

TEACHING APPROACHES

Having presented a framework one must consider what teaching approaches are suitable. Three basic principles are important [8 p.41].

1. Concepts and techniques will be developed in a practical context.
2. Techniques will not necessarily be completely developed at the time they first arise. Many of the ideas first introduced in the early years will continue to arise in later years.
3. A full theoretical justification of all the topics is neither necessary nor desirable. Some items will only be touched upon in particular problems; other concepts will be covered only through experiments and will not be justified theoretically.

A mixture of teaching strategies needs to be used: exposition, discussion, investigation, project work, simulation, case studies and practical work. The starting point should always be a problem as statistics is used to solve problems. After that one should proceed using appropriate strategies to help children find their own solutions to the problem.

Exposition can play a role in this, indeed this has always been a fundamental ingredient of work in the classroom. But there needs to be a dialogue whereby teachers should take account of and respond to the answers which pupils give, even if these answers are wrong or unexpected. Indeed statistics is the subject of uncertainty, making judgements in this situation is crucial. This can lead to worthwhile discussion and increased awareness for both teacher and pupil.

It is also useful for pupils to discuss ideas themselves, perhaps in small groups. This can be particularly valuable in a brain-storming session if a problem seems particularly intractable. Discussion can also clear up misunderstandings or help to clarify ideas.

The idea of investigation is fundamental. Investigations can be quite substantial and lead into individual project work, but this need not be the case. Indeed, perhaps more commonly, an investigation need be neither lengthy nor difficult, perhaps arising in response to a pupils' question such as "what would happen if...". Very often the question may be resolved quite quickly, sometimes it will need to be discussed at another time. The essential requirement is that pupils should be encouraged to think in this way and that the teacher should utilise such opportunities. There should be willingness even to follow some false trails.

Simulations can play a key role in presenting relatively difficult ideas in an intuitive way. Sometimes a full theoretical analysis may be rather difficult or even impossible but a simulation can throw light on the underlying situation and help pupils build up statistical intuition. This is particularly important in tackling probabilistic notions some of which can be quite tricky. The key idea is to simulate the uncertainty inherent in a particular situation and hence develop a deeper understanding. The advent of microcomputers will revolutionise the place of simulations as a large number of trials can be carried out effectively, reducing the drudgery. More details about simulations and Monte Carlo methods can be found in the NCTM yearbook [5].

Practical work is essential in putting over the main ideas behind data collection. This is a far wider and diverse area than in the traditional approach which pays only lip service to data collection before moving into often meaningless calculations. For, if one ignores data, there is little purpose in performing calculations of the mean, median or standard deviation. Indeed it is only by looking at data carefully that one can decide which measures are important. One also needs to be aware of the problems inherent in data collection - inaccuracies, omissions, bias. Children can only learn about this by trying to collect data in practical situations. This can be done by a small census, a questionnaire or a survey; sometimes one might use published data but one needs to be aware of appropriate sources. Occasionally one might only be able to take a sample, in which case care over inferences is necessary. In fact sampling is widely used in society and children need to be aware of the power and limitations of such techniques, preferably by taking samples themselves. Thus practical work lies at the heart of statistics, particularly in building up a firm background.

To facilitate effective teaching some equipment and resources are necessary but the essentials are relatively inexpensive [Holmes, 1981]. Micro-computers are likely to play an important role in statistics teaching but they would probably be bought as a resource for the whole school; besides their price is falling rapidly.

To summarize, it is important that all these techniques are used in the teaching of statistics. A problem-solving approach is essential as a starting point after which a variety of strategies should be used as appropriate. An example is presented below to give an idea of what is envisaged.

Retail Price Index

The retail price index is one of the major measures of the national economy now that inflation seems here to stay. This affects everyone yet few people understand how it works.

A suitable starting point for children is their weekly pocket-money. One might ask whether pocket-money this year has the same purchasing power as last year. This will lead into an interesting discussion of price rises of items such as sweets, crisps or drinks. Pupils can then try to construct their own individual little index using just four items by comparing present prices with those of the previous year. There is scope here for discussing accuracy of data and choice of items for the index. A pie chart or some other form of representation can be drawn. One can then go on to the idea of weights; clearly one buys more of certain items than others. Each pupil can choose his/her own items and weights to calculate a personal index number. This allows comparisons to be made within the class and helps to emphasize that the retail price index cannot reflect the effect of price rises on any individual but is a combined average. An appropriate average can be calculated from the class results.

After this it is appropriate to look at grocery prices in more detail, perhaps including a survey in local shops. Each pupil can find out what his/her family buys each week to build up a sort of Family Expenditure Survey. Thus one can work out a food index in a way which mirrors what is actually done in practice.

Finally one can broaden the index to include other commodities. The ideas of a clothing index and the forms that could take for families of different sizes can be considered. This leads naturally into an explanation of how the Retail Price Index is actually worked out, using the relevant published figures.

This teaching outline shows how various strategies can be used within a problem-solving framework. There are opportunities for exposition by the teacher, discussion between pupils, simulation, investigation and practical work. Techniques covered include problems of obtaining data, data tabulation and representation, applying measures of location and dispersion, index numbers, and comparing data. Clearly pupils will not become proficient in all these areas simultaneously but, by using this approach, they will meet these techniques in an appropriate context and, hopefully, learn how to apply the techniques in other situations.

Extensive evaluation of this approach has been conducted by the author as part of the research on the Schools Council Project on Statistical Education. These ideas were developed in a set of 27 units of work which have been carefully tested in a wide range of schools with teachers from varying backgrounds, teaching classes of different abilities. In general the materials produced were successfully taught to pupils of most abilities. More details can be found in Kapadia (1980). The materials have now been published [7].

CONCLUSIONS

Statistics is a hard subject to teach properly. By its very nature it is uncertain. In conducting experiments or surveys and making inferences one cannot predict what might happen. The statistics teacher needs to be particularly alert. Statistics also involves open-ended questions which have no 'right' answer. This is why a problem-solving approach linking techniques to applications is vital.

Probability also poses problems, not least because of the counter-intuitive nature of some of its results (do you believe the birthday paradox?). Another aspect which is hard to put across is the way that probability acts as a model in many situations. The fundamental point of modelling a supposedly unique event, like death, with probability is a difficult idea to accept. One knows it works (else insurance companies would have become bankrupt long ago) yet there is still a feeling of unease. It is by exposure to such ideas through simulation and investigation that children will learn to come to terms with these concepts. To facilitate understanding one should start with a frequentist approach estimating probabilities from experimental data; this emphasizes the link with statistics. It is also important to develop a subjective approach as well as using a priori considerations. These ideas are only presented briefly as this paper is mainly concerned with statistics. Yet the theoretical basis afforded by probability must not be ignored.

This paper has presented the outline of a course suitable to provide all children with a necessary background in statistics. The historical perspective and modern applications provide the rationale. The aims of the course must be to link objectives and applications within a problem-solving framework. A variety of teaching approaches is necessary as illustrated by the example of the retail price index. The general overall aim is to achieve a level of statistical numeracy so that all pupils learn to appreciate the powers and limitations of statistics.

References

[1] Cockcroft, W. (1982) Mathematics Counts, HMSO.

[2] Hacking, I. (1981) How should we do the history of statistics? I. & C. No.8, pp.15-26.

[3] Holmes, P. et al (1981) Statistics in Schools 11-16: A Review, Methuen.

[4] Kapadia, R. (1980) Developments in statistical education. Educ.Stud.Maths., 11, pp.443-461.

[5] Shulte, A. (ed) (1981) Teaching Statistics and Probability, N.C.T.M.

[6] Pearson, E. and Kendall, M. (eds) (1970) Studies in the history of statistics and probability, Griffin.

[7] School Council Project on Statistical Education (1981) Statistics in Your World (27 pupil booklets, 27 teacher booklets), Foulsham.

[8] School Council Project on Statistical Education (1981) Teaching Statistics 11-16, Foulsham.

[9] Tanur, J. (ed) (1972) Statistics: A guide to the unknown, Holden-Day.

ON THE TEACHING OF ORDER STATISTICS

K Bognár, Eötvös Loránd University

To deal with order statistics is very fruitful from several aspects. Besides getting the pupils acquainted with different statistical notions (median, range, etc.), we have an opportunity to show how to distinguish between independence and dependence and the different levels of dependence (correlation). The variety of combinatorial methods used in connection with order statistics is also an argument in favour of teaching this topic.

The program to be presented is for pupils of 13-15 years who learned some stochastics (relative frequency, probability, mean etc.) within the new mathematical curriculum.

The problems were usually connected with games, which has several advantages. The collection of data (which is essential in the teaching of statistics) is not so monotonous if an experiment is performed for deciding who is the winner, whose tip is good. The games are generally such that pupils may choose from different strategies and this gives them some means of "defending" themselves against chance.

The games were played in the classroom with coins, balls, dice and pearls but, later on, pupils used tables of random numbers. By the aid of the usual tables of random digits they prepared tables of 0-1 sequences in which the relative frequency of zero is not $\frac{1}{2}$.

GAMES WITH ZERO-ONE SEQUENCES

One may disagree with treating 0-1 sequences in connection with order statistics, since many facts become trivial in this case but - in our opinion - 0-1 sequences help to understand notions just because of their simple character.

Instead of tossing a fair coin (or unfair coin), we draw (with replacement) from a box containing a "zero" ball and a "one" ball (or some "zero" balls and some "one" balls). Thus we generate 0-1 sequences, say, of length n. The probability of drawing a zero will be denoted by p_0.

The digits generated will first be written as they come out (unordered sample) and then they will be ordered nondecreasingly (ordered sample).

GAME 1:

n will be fixed as 3, 4, 5; $p_o = 0.5$. The player writes two 0-1 sequences of length n: a tip for the unordered sample and one for the ordered sample. Under each digit in their tips they write a nonnegative integer (stake), the sum of these integers in each tip should be 2n. Also they pay 2n toplak to the bank. After that a 0-1 sequence is generated, and for each hit the players get the amount appearing under the corresponding digit in their tip.

The game can be played either by the whole class together or in pairs; in the latter case much more random sequences will be obtained. Pupils prepare the following table:

	unordered sample								ordered sample						
tip	1	1	1	1	1			pure gain -1	0	0	1	1	1		pure gain 5
stakes	2	2	2	2	2				10	0	0	0	0		
generated sequence	1	0	0	1	0			number of hits 2	0	0	0	1	1		number of hits 4
hits	X			X					X	X		X	X		
tip...															

Questions to be discussed:

1. Are there preferable strategies in case of unordered and in case of ordered samples?
2. Which n ($3 \leq n \leq 10$) do you prefer?
3. Which p_o do you prefer?

GAME 2a:

First let $p_o = 0.5$. Before generating 0-1 sequences of length 5 pupils write (secretly) their tip concerning the last zero in the ordered sample. (If they think there will be no zero in the sequence, they write O). Those whose tip is correct get 5 toplak; those who err by one get 4, those who were mistaken by 2 get 1 toplak; the others get nothing.

GAME 2b:

The same game, but now the 0-1 sequences are generated with the aid of a box with 2 "zeros" and 1 "one" ($p_o = \frac{2}{3}$).

To the question "How many toplak are you ready to give for a game?", most of the answers were 3 (which is not bad, since the game is profitable up to 3.79 when the player's tip is 3).

The pupils determined the distribution of $X_i^{(5)}$ (where $X_i^{(n)}$ denotes the i-th digit of the ordered sample of length n) and compared it with the empirical distribution (see Table 1, $p_o = \frac{2}{3}$).

i	$P(X_i^{(5)} = 0)$	Empirical (200 sequences)
1	0.996	0.99
2	0.955	0.96
3	0.790	0.77
4	0.461	0.46
5	0.132	0.12

TABLE 1

Afterwards, it was easy to discuss the fact that it is a bit easier to find out the ordered sample in case $p_o = \frac{2}{3}$ than in case $p_o = \frac{1}{2}$.

GAME 2c:

A special case of the Polya model is considered. At the beginning we have two balls in a box, numbered 0 and 1 again. A ball is drawn at random from the box, its number is written down, then (before the next draw) it is replaced and also an additional ball of the same type is put into the box. This process is repeated five times. Thus we get as a result a Polya-type 0-1 sequence. The tip concerns again the last zero in the ordered sample.

Pupils can see after a short time that these 0-1 sequences behave differently from the sequences generated by a simple replacement. The empirical distribution of the last zero (see Table 2) was surprising to them.

i	0	1	2	3	4	5
Empirical (150 sequences)	0.15	0.21	0.15	0.17	0.15	0.17

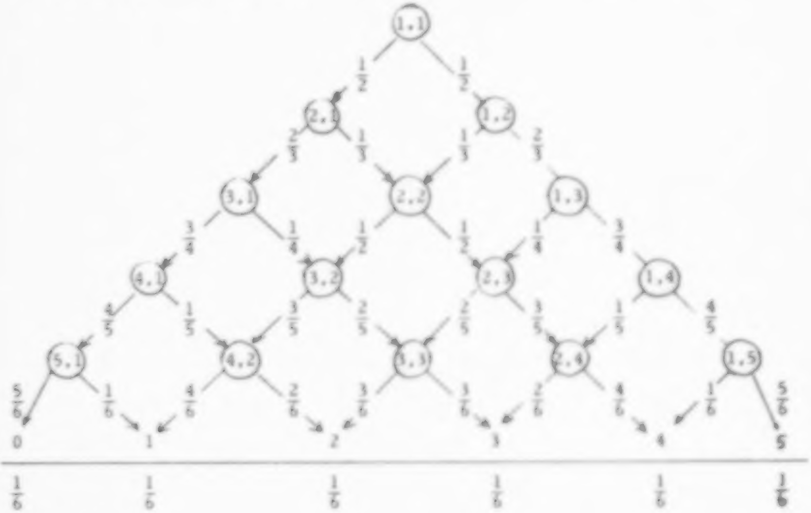
TABLE 2

Observing always say, the second digits in the unordered samples, pupils could see that nearly half of them were zero (as if they were drawn with replacement from a box with $p_o = \frac{1}{2}$). Further, pupils could see that the connections between first and second; first and third; first and fifth digits in the ordered samples became looser and looser.

GAME 3:

Player A tells an integer n, this will be the length of the 0-1 sequence to be generated ($p_o = \frac{1}{2}$). Then B tells k ($1 \leq k \leq n$) which means that A should find out the k-th digit of the ordered sample. A gets (n-k) toplak from B if A's tip is good, and A pays 2(n-k) toplak (or k(n-k) toplak in another version) in case of a mistake. The average gains of A (by different strategies) can be seen in Table 3 (or in Table 4 in case k(n-k) toplak). The pupils did not compute these, but they found out good strategies. Pupils were asked to draw a scheme and indicate their gains by jetons of two colours (positive and negative), from which everybody could see which strategies were good for A and B. They soon discovered the difference between n=2s and n=2s+1.

This game was played also with Polya-type sequences. Pupils evaluated the exact distribution of the number of zeros with the help of the following graph (zero to the left, one to the right).



GAME 3a:

A version of Game 3 is the following. Player A tells an integer n, then B tells i, k ($1 \leq i < k \leq n$), which means that A should find out the i-th and k-th digits of the ordered sample of a random 0-1 sequence of length n. A gets 3 toplak from B for a correct tip and pays to B 2 toplak in case of any mistake.

To find out how to play well, the pupils evaluated the joint distribution of the i-th and k-th digit of the ordered sample.

B \ A	2	3	4	5	6	7	8	9	10
1	0.3	1.3	2.4	3.6	4.8	5.9	6.9	7.5	9.0
2	0	-0.5	0.1	1.3	2.7	4.1	5.4	6.6	7.7
3		0	0.1	-1.0	-0.1	1.3	2.8	4.4	5.9
			0	0.4	-0.1	-1.5	-0.4	1.2	2.9
5				0	0.7	0.6	-0.3	-2.0	-0.7
6					0	0.8	1.1	0.7	-0.5
7						0	0.9	1.5	1.5
8							0	0.9	1.7
9								0	1.0
10									0

TABLE 3

B \ A	2	3	4	5	6	7	8	9	10
1	0.5	1.5	2.6	3.6	4.8	5.9	6.9	8.0	9.0
2	0	-0.5	0.1	1.3	2.7	4.1	5.4	6.6	7.7
3		0	-0.3	-2.0	-1.1	0.4	2.1	3.8	5.5
4			0	0.1	-1.4	-4.5	-3.3	-1.3	0.8
5				0	0.3	-0.7	-3.5	-8.0	-6.3
6					0	0.6	-0.02	-2.3	-6.6
7						0	0.7	0.6	-1.1
8							0	0.8	1.0
9								0	0.9
10									0

TABLE 4

GAME 4:

Each of two players A and B generates a 0-1 sequence ($p_o = 0.5$) until at least one of them obtains a "run" 11 (i.e. the digit 1 appears twice one after the other).

1. If both A and B obtained their first run within 10 trials, namely k_A trials needed for A and k_B for B, then

a) in case $k_A = k_B$, the game is finished and nobody pays

b) in case $k_A < k_B$, either B pays $k_B - k_A$ toplak to A, or B begins to generate a new sequence (in which case A may also choose to begin a new sequence) and the rules are applied for the last sequence of each player.

2. If A obtained while B did not obtain a "run" within 10 trials, then B either stops and pays $10-k_A$ to A or generates a new sequence until a run appears and pays the length of this new sequence to A.

3. If neither A nor B obtained a run within 10 trials, each of them should generate a new sequence and they pay according to 1.

Denote by P_i the probability that the number of drawings needed until a run 11 appears is just i. Table 5 contains these probabilities for a few i's (evaluated by the pupils).

i	2	3	4	5	6	7	8	9	10
P_i	0.250	0.125	0.125	0.094	0.078	0.063	0.051	0.041	0.03

TABLE 5

GAMES WITH ORDERED SAMPLES FROM UNIFORM DISTRIBUTIONS

To get nearer to more realistic situations we treat ordered samples from the uniform distribution on $0,1,\ldots,r$ $(r > 1)$.

GAME 5:

A fair die (numbered from 0 to 5) will be thrown in this game 10 times. Before that, everybody writes 10 integers (as tips for the results), and orders The generated sequence is also ordered, and is written under the tip sequence. The differences between the corresponding digits of the tip and the generated sequence are computed, and +1, -1 or 0 is written in the corresponding place, according as the difference is positive, negative or zero. Thus everybody has a sequence of elements +1, -1 and 0 (error sequence). If a "random walk" (starting from the origin) with this error sequence remains in the interval [-3, 3], the player gets 2 toplak; if in [-2, 2] then 5 toplak, and if in [-1, 1], then 10 toplak.

Pupils listed the maximal distance from the original (M) and also the sum of the digits in the error sequence (S) (Table 6).

Tip	1	1	2	2	3	3	4	4	6	6			
Ordered Sample	1	1	1	2	3	4	4	5	6	6	M	S	Gain
Error Sequence	0	0	+1	0	0	-1	0	-1	0	0	1	-1	10

TABLE 6

GAME 6:

The bank has a box containing x balls numbered from 0 to x. The player should find out x by drawing with replacement from the box 3, 4 or 5 times. The player should pay the bank 1 toplak for each drawing and gets from the bank x, x-1 or x-2 toplak respectively if he hits x, errs by one or two respectively; otherwise he gets nothing (x may be between 3 and 10).

The average gains of the player (when estimating x by the maximal value of the numbers drawn) can be seen in Table 7.

Average gains of the player in Game 6

Number of draws x	3	4	5
3	-0.6	-1.4	-2.3
4	0.1	-0.6	-1.4
5	0.7	0.1	-0.6
6	1.3	0.8	0.1
7	1.7	1.3	0.8
8	2.0	1.8	1.4
9	2.4	2.3	1.9
10	2.7	2.6	2.4

TABLE 7

GAME 7:

A box contains r balls numbered from 0 to r. From this box n balls are drawn with replacement. The range of the sample (the difference between the largest and the smallest digit) is computed. Players write a tip for the range before generating the sequence and get 5, 4 or 1 toplak from the bank according as the absolute value of the difference between the tip and the actual range is 0, 1 or 2 (otherwise they get nothing).

Pupils evaluated the exact distribution of the range for r=5, n=3; for r=5, n=5; for r=9, n=5 and decided which is more profitable and which sum is worth paying to the bank for a game.

For r=3 (die) they examined the connections between digits of different distances. For n=3 they had a lot of empirical data from which they evaluated the conditional expectation of the i-th digit given the value of the first digit. Table 8 contains their results.

Value of the first digit	Conditional expectation of the			
	second digit		third digit	
	Theoretical	Empirical	Theoretical	Empirical
0	1.81	1.78	3.63	3.72
1	2.48	2.53	3.95	4.10
2	3.14	3.26	4.27	4.30
3	3.79	3.77	4.58	4.56
4	4.43	4.24	4.86	4.88
5	5	5	5	5

TABLE 8

They offered measures for the connection with the aid of this table. They offered games in which players buy information concerning the value of a certain digit.

GAME 8:

Player A puts (in secret) some white and some red balls (at least one of each colour) into a box, so that the total number of balls in the box is 10. Another box is prepared by A with the same content, both given to B. Then A tells B that there are either k_1 or k_2 red balls (among 10) in each box (the same number in both). B should find out which in fact is the case. B draws with replacement simultaneously from both boxes till a red ball comes at least from one of the boxes. If the last balls both happen to be red, then B is obliged to stop. Otherwise B may stop. In both cases of stopping B makes a decision (between k_1 and k_2). B pays one toplak to A for each (double) drawing. Further, B gets (pays) $\frac{10}{|k_1-k_2|}$ from (to) A if the decision is correct (incorrect). If B does not stop, he may go on drawing from the other box till a red ball appears there also (but no longer). The sums to be payed are as above.

Denote by $X_1(k)$ and $X_2(k)$ the number of drawings needed till a red ball appears from the first and second box respectively (each containing k red balls). Further denote $q = 1-\frac{k}{10}$.

Table 9 contains some values of the probability distribution of minimum $(X_1(k), X_2(k))$. $(P(\min(X_1,X_2) = i) = q^{2(i-1)}-q^{2i})$.

i \ q	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9
1	0.99	0.96	0.91	0.84	0.75	0.64	0.51	0.36	0.19
2	0.01	0.04	0.08	0.13	0.19	0.23	0.25	0.23	0.15
3	0.00	0.00	0.01	0.02	0.05	0.08	0.12	0.15	0.12
4	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.09	0.10
5	0.00	0.00	0.00	0.00	0.00	0.01	0.03	0.06	0.08
...									
	$P(X_1(k) = X_2(k) = i)$								
1	0.81	0.64	0.49	0.36	0.25	0.16	0.09	0.04	0.01
2	0.01	0.03	0.04	0.06	0.06	0.06	0.04	0.03	0.01
3	0.00	0.00	0.00	0.01	0.02	0.02	0.02	0.02	0.01
...									

TABLE 9

References

[1] Engel, A. (1976) Wahrscheinlichkeitsrechnung und Statistik, I-II. Klett: Stuttgart.

[2] Rényi, A. (1969) On the teaching of probability, Rouse Ball Conference.

[3] Rényi, A. (1970) Foundations of probability. Holden-Day, Inc., San Francisco, Cambridge, London, Amsterdam.

[4] Vincze, I. Introduction into the theory of order statistics... (Manuscript).

[5] Bognár, K. (1980) Random fluctuation through games. "Stochastik im Schulunterricht". Hölder-Pichler-Tempsky, Wien, B.G. Teubner, Stuttgart.

[6] Bognár, K., Nemetz, T. and Tusnády, G. (1980) Ismerkedés a véletlennel. Tankönyvkiadó, Budapest (In Hungarian).

TOPIC 3: TEACHING STATISTICS IN SCHOOLS
(16-18 YEARS AGE GROUP)

ORGANISER: PROFESSOR J SWIFT

AMOUNT AND KIND OF STATISTICS TO BE TAUGHT TO 16+ YEAR OLDS

F Eicker, University of Dortmund

PRESENT STATUS OF STATISTICS AT SCHOOLS. FUTURE STEPS.

The lectures at this conference have shown once again a broad consensus among statisticians on arguments in favour of the school subject statistics and also on general issues of how to teach it: e.g. the importance of education in numeracy in today's world, the teaching should start at a young age, realistic data of interest to the pupil exist abundantly and can be used at school. But also quite a few details of the teaching methods and contents seem to find widespread approval: e.g. certain proposals on how to introduce and to apply specific concepts or methods, how to tie probability to statistics, to let children collect their own data, etc.

Our future work on teaching statistics and at conferences such as this is likely to concentrate on balancing the new subject of statistics with the other school subjects, on further details of instruction and particularly on sharing internationally the teaching experience already available. This procedure may lead rather quickly to a sufficiently large collection of tried out teaching topics from which the (modest) content of a statistics syllabus may be chosen to suit the needs of the particular groups of students at hand. Many problems are similar internationally as again this conference has shown. New textbooks may then be developed like 'Statistics in Your World' for the 11-16 year olds [5] and existing textbooks may be adapted. When good teaching material exists which represents well enough the statistical activities this will provide the basis for a new discussion of the nature and of the place of statistics at schools. This certainly is necessary in many countries, e.g. in Germany.

Since I am not a school teacher I cannot contribute anything regarding classroom experience. But since teaching 16-19 year olds and teaching statistics to non-statisticians at college have much in common, let me describe some personal experiences and ideas, and let me propose a few methodological details and sources of applications useful at schools.

But let me first, though, recall some characteristics of statistical work in general and some of its implementations for school teaching. Let us reflect in particular how it differs from mathematics. All this of course is basic to our efforts. It is not new what I say here but it seems necessary in many countries, particularly in mine, to draw the attention to the very nature of our subject in order to present the field of statistics appropriately in our teaching and to guard against unforeseen developments such as they occurred e.g. in set theory at schools.

THE NATURE OF STATISTICAL REASONING

Let me illustrate by an example some relevant aspects of the nature of statistical reasoning. Part of the example seems also directly suitable for school instruction: let us look at regression analysis which is probably the most widely used and also a most typical statistical technique (for a typical recent book see [1]). Its general aim is to try to represent observed values of a variable y sufficiently closely as a certain function of the values of another observed variable x usually called regressor or predictor. If nothing is known about the causal relationship between both (e.g. between height x and weight y of a growing plant) one would of course better look for a more general interrelation $g(x,y) = 0$ between both. (The case of two or more regressors is frequent in practice but needs not be taken up at schools nor for our present purpose).

The closeness of fit of f depends on the purpose of the investigation. Often neither the type of the function f nor the degree of closeness of the fit are known in advance but have to be tentatively or successively determined. Typically this is a problem of data analysis or of descriptive statistics. The representation can be considered a mathematical model $y = f(x)$ in case the causal relation between x and y is clear (as e.g. between the weight of y of a plant and the amount x of water or a fertilizer). Then also the name predictor for x becomes evident. The simplest case here is a linear function f of only one regressor: $y = a + bx + e$ which can always be assumed in case both the x and the y-values are clumped in small neighbourhoods and the global function f is smooth on the x-neighbourhood.

Any relevant prior information from the empirical object of study (which has nothing to do with Bayes' prior distribution) regarding the structure of the function f, the kind and number of regressors, the errors e etc. should be exploited in this whole process. The previous knowledge largely belongs to the empirical sphere of the problem but clearly must be investigated and understood by the statistician. This part of his responsibility thus has the nature of an inductive investigation typical for quantitative empirical science. If no or only insufficient prior knowledge is available it may be substituted by structures obtained from exploratory data analysis as in the above regression example indicated.

Once these steps have led to a mathematical model the derivation of its mathematical properties of interest has the nature of purely deductive reasoning and thus is part of mathematics, in particular often of probability theory and mathematical statistics. But the reinterpretation of deduced results into the empirical context then again leads out of the domain of mathematics.

Thus statistical analysis comprises both induction and deduction simultaneously and consciously or not, and regression is very well suited to exemplify this according to the very brief description above. Regression is also very well suited under didactical aspects since there are very simple as well as very complicated situations and an abundance of interesting examples.

An important conclusion we derive from this and many similarly suited examples of statistical analyses is that statistics can properly be subsumed neither under mathematics nor under any empirical field alone. If we could reach general agreement on this it would answer many futile disputes once and for all. Moreover, the relevant course of the school curriculum should appropriately be named just statistics or at least 'mathematics and statistics' instead of mathematics alone.

It must be emphasized that the empirical work includes all relevant technicalities of the often very complicated data gathering processes. These should be subject to statistical requirements or control whenever possible (e.g. in controlled experiments). Consequently the statistician must cooperate most closely with the empirical worker who as a rule does the actual data collection.

The above account suggests that statistics can be defined as a methodology to acquire quantitative knowledge under uncertainty. (The definition occasionally made as decision making under uncertainty is too narrow). Its duality of approaches leads to calling it an 'art and science'. Remarks to this end can be found in many forewords of statistical textbooks, in the International Encyclopaedia of Statistics [3] under the article on Statistics by W. Kruskal and elsewhere.

It should still be mentioned that often only higher precision of the results requires the use of probability models and tools in statistical analysis, unless prior information demands a specific probabilistic model. Since in many cases and in particular at school high precision is not required, purely data analytic methods constitute an important branch of statistics and of teaching it. They are often mathematically fairly easy, so they can be treated conveniently at school. Since in hardly any cases where precision is needed the probabilistic model assumptions are acceptable beyond any doubt, data analysis is still an indispensable tool in establishing or checking these assumptions (e.g. checking residual plots by eye). Thus statistics instruction at school restricted to descriptive data analysis may often be completely sufficient. Part of it can be taught at a very young age, other parts can be quite demanding mathematically (details and examples below).

Another general comment is that no misleading illusion about the potentials or the scope of statistical methods should be awakened (e.g. if the data base is shaky). Some of the limitations can be inferred from the above said. On the other hand it should not be reaffirmed statistically what is evident beyond any doubt simply by looking at the data.

In summary then regarding teaching statistics at school it should be clear from the above that

- the didactical and organizational problems are quite specific and should be well contemplated;
- the limitation to the mathematical aspects or foundations only is no substitute for a proper statistics course and would undoubtedly lead to a quick end to the episode of statistics instruction at schools;
- a balanced presentation of statistics can provide lots of excellent motivation for any age and ability age group of pupils. It thus can foster objective and mathematical thinking for all.
- overselling statistics at school should be avoided as carefully as underselling it.

TARGET GROUPS

A discussion of the amount and the kind of statistics to be taught at school in any one country must differentiate among the various target groups of students to be instructed. A possible statistics programme for those leaving school at 16 will look quite different from that for students up to 18 or 19 years. And again for the 16+ group there will be different programmes for those concentrating on vocational or on general education courses. We shall subsequently not discuss specific requirements for special types of schools or course programmes but rather concentrate on general education courses. To a large degree, although this varies temporally and geographically, these students will not study at universities later on, and again, of those who do go on to universities, only a fraction will ever use statistics professionally or even study mathematics, science or statistics. Hence it certainly is not justifiable to adjust statistics programmes to the specific needs of some minority groups. The pedagogically much more demanding task and the greatest need rather arises from the group of students that later on uses statistics peripherally in business and industry, in the biosciences and health, sociosciences, psychology and other areas. Often enough those taking jobs in these areas may actually get their only statistics education at school. But an equally important challenge lies in the above mentioned citizen awareness type of education in an age when many published statements use (and abuse) statistical concepts. Still another need exists in the other school subjects taught concurrently. In summary then it should be clear that we must concentrate our discussion on the modest needs of an average user (or consumer) in the above sense and from this there arises a didactical task quite specific for the field as should have become clear by delineating the very nature of statistical reasoning.

Since a statistics syllabus in most countries has not yet been rigidly introduced, much of this article concerns both the pre-16 and post-16 age group. Where appropriate a dividing line is pointed out.

SELECTION OF STATISTICAL SCHOOL TOPICS

The case of regression mentioned earlier is but one example from the wide range of statistical subjects. We have to choose very few representative topics for school, perhaps one or two: regression may be a possible one because of its flexibility. Obviously there is no optimal choice. Nor do we want to train a mini-statistician. Rather it would be a serious question what a minimum amount of teaching material is that still makes sense. But we are not trying to answer it. Certainly the whole task is very much eased if elementary concepts and examples are already taught early during schooling. This possibility certainly exists and is well tried out (cf. the recent school book series 'Statistics in Your World'). A separate session as you know is devoted to this question. Let me now loosely list and discuss some themes which may be candidates for school. Of course the list is incomplete and I may have forgotten even some very important topics. In the case of each and every subsequently suggested method this method should be strictly tied into a real life motivating problem. This is usually a time-consuming job to teach.

To begin with the children should collect their own data a number of times. If this is done at a higher age (e.g. to take interviews) the students may in addition describe facts that influence or determine a probability model for the sampling process. Elementary data analysis can be done early. Location and scale can be discussed and investigated via median or mean, resp. range and variance. These measures must not just be calculated and then left but they must be interpreted in the empirical context. For higher grades more complicated methods of exploratory data analysis may well be up to the mathematical abilities, e.g. regressions, fitting curves by least squares, smoothing, trends, seasonal variations, even lagged variables, forecasting, orthogonal regressions and the question of causality, possibly also rank correlation. Even a limitation only to examples from descriptive statistics might seem reasonable. It would avoid the lengthy and delicate introduction of probabilistic inference, model building etc. which would almost require a separate course if done properly. But such a restriction should not rule out what might be called 'intuitive inference' (G. Noether). Descriptive statistics must always even at the earliest stages provoke the question 'What do these figures tell me?'. It is almost needless to say that they can only be taught by way of real life examples. It is anyway the natural preparation for probabilistic inference: data analysis often gives hints for particular probability models and it is used to check their validity and the inferences based on them. We said above that this is the way one proceeds in statistical practice. The nature and value of inference statements can then be better understood with their help. There is a lot of material illustrating this from statistical practice and much can be used in school.

Only a few of the simplest examples of inferential statistics should be explained. These examples are often excellent ways of showing the value of 'means' and similar functionals. In particular discrete, finite valued random variables and distributions suffice. But the real meaning and validity of (simple) stochastic models should be deeply and critically discussed and understood (if discussed at all). Elementary non-parametric methods are superbly appropriate in this connection, as has been advocated for a long time e.g. by G. Noether. Point estimation and confidence interval estimation should be preferred to testing.

Probabilistic tools should be kept to the utmost minimum and should only be introduced if their bearing on statistical applications is clearly and immediately visible to the students. Actually it has often been advocated by teachers and statisticians that at the school level all that is needed is an intuitive understanding of the frequency aspect of probability. Much can be referred to tables or simulation experiments. But if probability is really used it should always be the aim to ensure that the student really understands the leading ideas and that he knows what they mean in the real world and in a statistical statement, e.g. the concepts of randomness, stochastic independence, probability as a measure of uncertainty etc.

It should be understood that this latter probability concept is actually often applied in everyday life to such non-frequentist applications as the probability of a horse winning a race, a particular football team winning, of passing a particular examination. But still this is not a reason for including a substantial course on probability theory for any age group, e.g. the impression must strictly be suppressed that the concept of an integral (or even improper integral) is a necessary prerequisite for handling simple statistical problems. On the other hand statistics and its application offer a wealth of opportunities to select and illustrate any number and any level of mathematical concepts and techniques that are possibly desired in general school instruction. It is thus in perfect agreement with a strong and lasting strand in formulating the goals of mathematics education in general as expressed e.g. in 'Studies in mathematics education' [4] (cf. in particular vol.2, p.159ff): it consists of stressing its relation to and bearing on many needs of real life problems.

By contrast a course on probability alone (or one with just a brief formal appendix on a topic of theoretical statistics) would be much more in the strand of abstract mathematics teaching. It would then also be in competition with other purely mathematical topics. Such a course would not incorporate the motivation by statistical applications. It must therefore be expected to have the fate of a brief transitional fashion in advanced mathematical school instruction like many other themes in the past and in the present. It would be fully in the tradition of most mathematics instruction. It also does not make a difference if such a course carries the new name 'stochastics' (as in Germany) which by definition excludes statistics in a comprehensive sense.

Continuing the survey of possible school topics it should be added that when the pupils actually take their sample also the defining questions and the practical problems connected with it should be pointed out to them, e.g. the question what is a household or what errors can occur when taking an interview. Experimental design may be mentioned, no matter whether the issue is spelled out in detail. Although usually the statistician does not himself collect the data he must know and often control everything relevant to his inference about this process. Hence also the students should hear about these steps.

Another good source of examples in sampling, random sampling or stratified, is the (rough) estimation of finite population parameters from such sampling. The probability model in this case is quite easy to obtain by the way.

Some statistical topics that actually occur in geography, economics, biology, science as taught today to the post-16 group are, according to findings of the Schools Council Project at Sheffield [2], the design and use of questionnaires (and interviews), reliability of data, index numbers, averages, trend analysis. Some, or all (and many more) of these basic concepts are vital, and they can be taught in statistics courses. Their importance as well as their presentation obviously need not or cannot be compared with traditional mathematical teaching topics. It is our didactical challenge to incorporate such new concepts into school instruction, to prove that this is possible and then press for an adaptation of the curricula.

Good sources of real life examples and often a sufficient description of methodology can typically be found e.g. in good textbooks on introductory applied statistics for biology, medical and economics students; they exist in many languages.

It also seems obvious today to use electronic computers in the actual statistical calculations. There are many encouraging experiences in this direction. For detail refer to proper sessions of this conference.

Whatever is going to happen, a self-critical and unambitious attitude is appropriate with regard to the total amount of teaching material as well as the degree of sophistication. A few enlightening and really well understood encounters with statistics within an empirical context in the pupil's entire school career seem completely sufficient and at the same time all that can realistically be hoped for. For instance in the 16+ age group a practical case or two of sampling and estimation of the population parameters from such sampling might fully suffice. On the other hand time and again students should be reminded of the existence of statistical investigations and methods applied to a wide variety of cases so that they can see the flexibility. But as much as an attempt at a systematic coverage we must avoid the mechanical reproduction of some cookbook recipes. Considering the long list of possibilities, demands and restrictions there is much hard work still to be done on practical teaching beyond the consensus on general issues.

One of these acute tasks is teacher training. There are again special sessions on this, so let us remark just one or two things. Obviously the difference between statistical and mathematical work must be effectively reflected in teacher training, possibly also in (new) university courses. The work and scope of the average teacher, i.e. of their majority should be taken into account. The development of realistic training programmes as well as of curricula will require patience and time. It is not easy to design the appropriate statistics training programme for teachers. For instance courses taught other than by mathematics departments may be appropriate and suitable for teacher students. But this problem must be realised and tackled or the experiment 'statistics in schools' will become another shortlived and disappointing episode in the history of school mathematics.

Acknowledgements

For critical comments and additions on a preliminary version of this paper the author is much indebted to G. Fillbrunn, P. Holmes, H. Kütting, T. Nemetz, G. Noether and J. Swift.

References

[1] Draper, N.R. and Smith, H. (1981) Applied regression analysis. 2nd ed. Wiley.

[2] Holmes, P. et al (1981) Statistics in Schools 11-16: A Review, Methuen Educational.

[3] Kruskal, W.H. (1978) Statistics. Internat.Encycl. of Stat., The Free Press, Div. of McMillan Co., New York.

[4] Morris, R., (ed) (1981) Studies in mathematics education, 2 vols., UNESCO, 7 Place de Fontenoy, 75700, Paris, France.

[5] Schools Council Project on Statistical Education (1981) Statistics in Your World, School book series and teacher's guide. Foulsham Educational, Slough,UK.

TEACHING STATISTICS TO 16-18 YEAR OLDS: DREAMS AND REALITY

T Nemetz, Mathematical Institute of the Academy, Budapest

SUMMARY

Many statisticians are convinced that statistics, as an independent and compulsory subject, should be included in secondary-level general education. There are different (sometimes quite ambitious) proposals for the subject matter. Our real possibilities seem, however, not to match them. In this paper we discuss some aspects of this confrontation. Our conclusion is, that to a certain extent, statistics can be dealt with within mathematics. Examples are given to show how one can embed statistical notions into traditionally taught mathematics. Still much effort has to be taken to improve conditions for a widespread introduction of an independent statistics course.

DREAMS

Likely most statisticians have had adventures into the area of "what should and could be taught" at school level. It is even more likely that all of us can turn up with beautiful ideas and programs after just short consideration. In what follows we try to summarize the typical first-thought-conceptions. This is mainly composed of (a bit biased) quotations (without references), mixed with my own first notions.

The starting point of my dream is a perfect harmony among the three main disciplines of the area called stochastics: the descriptive, the mathematical statistics and the probability theory. The best men in these areas are eager to cooperate in finding the best possible answer to the question: what, at which level and how should it be taught. The educational authorities, as well as the teachers of all subjects are cooperative, too.

Since there is hardly any historical or didactical reason for separating these three branches, their teaching goes parallel, following a kind of spiral structure.

The goal of teaching statistics at schools is making the children familiar with statistical thinking with emphasis on influencing their thinking rather than providing them with recipes for specific situations. The subject matter aims at the "average" (in many senses) pupils and is assembled to provide them with a proper understanding of the surrounding environment and to satisfy their curiosities about the "behaviour" of a random media, which is vital for intellectual growth. There are many illustrations for the wide applicability of stochastic methods with proper examples from various fields. Such examples are also incorporated into the teaching of the basic disciplines like geography, biology, physics, economics etc.

An important aspect is that the pupils start getting acquainted with the world of randomness at a time when they are not yet infected with the notion of determinism, i.e., from the very beginning of the primary school. Statistical and probabilistic ideas are spread through the whole curriculum, from grade 1 onwards, allowing a gradual discovery of the "regularities in the uncertain". At the last stage a small, compact course is offered, containing formal, correct definitions and some theorems with proofs. Of course, this is an independent subject, and not a (possibly hidden) part of other subjects.

As to the goals and guidelines the dreams of many of us go more or less along the same lines. As to the details, however, the "first-thought-conceptions" vary from one extreme to the other. We bring just a few such extremes:

- Descriptive statistics is too specific to be taught. It suffices to mention just a few characteristics like the average, in connection with arithmetic.
- Descriptive statistics plays the most important role in everyday life, therefore its detailed teaching is of vital importance.
- Mathematical statistics is to be handled just to the extent of helping the teaching of probability (if included in the subject matter at all).
- I have encountered the proposal to include in the curriculum even the theoretical background of the method of chi-square tests.
- The children do not gain anything from the over-idealized/theoreticized area of the probability theory - no need for it.
- A real understanding of the stochastic ideas supposes the knowledge of the central limit theorems, therefore one has to insist on its teaching.

Luckily, most dreams are much more realistic. These include the dreams of those who had taken the trouble to work for a concrete teaching project. Hopefully this conference helps us to collect a short description of their last dreams.

REALITY

Below we list some of the main difficulties in realizing not only our nice dream, but even more realistic conceptions. Of course, our experiences in this respect are strongly influenced by the particular Hungarian conditions, where we had, till last year, a very rigid school system (see [3,4]).

Let us start with a mild disappointment: there seems to be a complete disharmony among descriptive statisticians, mathematical statisticians, researchers in probability theory and the ones who apply it. This disharmony leads to contrasting (if not conflicting) recommendations toward the educational authorities, who have by now accepted the idea of introducing statistics/stochastics in the curricula of the pre-university schools. Introducing statistics means a struggle for the authorities from the point of assigning time to its teaching. At present, this can be done to the detriment of some well-established subject only. And we have to admit that good teaching of statistics uses much time, especially at secondary level, and cannot be done in just a few lessons. And the authorities face the most serious problem: the lack of teachers with proper education in statistics.

Round table conferences on the teaching of statistics during ISI meetings show that the "teacher problem" is the biggest obstacle in introducing statistics to schools, all the world over. The initial training of teachers seldom contains even passing contact with statistics. Secondary level mathematics teachers have had some training in probability, but usually in a rather formal way, which does not help them to understand statistical ideas. Re-training teachers seems to be very difficult. Besides financial problems, the administration of re-training also makes difficulties, and the effectiveness thereof is questionable. We completely share the Polish experiences (W. Szlenk [2], pp.50-54): "... it is hardly possible for an adult to assimilate the elements... in just a two-week course... it is a very hard question how to teach an over 40-years old teacher. In my experience only a few of them could get used to notions like probability and independent events".

Some people have the opinion that this problem can be overcome by relying on teachers of other disciplines like economics, biology, social sciences etc. (see [2], pp.17,18). In Hungary this is no solution. Economics is taught in a few highly specialized secondary schools, and we are concerned with the education of "average" students in general secondary schools. Teachers of other subjects are lacking even the formal acquaintance with probability of the mathematics teachers. We can see some attempts in biology: among the 4500 test-questions assembled as candidate questions to the final examination in biology, there are 3 of statistical nature, one of which is given in the Appendix, Example 1, for judgement.

The examination-oriented teaching raises another difficulty. We have an "almost uniform" final examination in secondary schools in the main disciplines, also entrance examinations to universities. Due to this uniform nature, no one can expect a test-question about some statistical problem in any subject while it is taught just in part of the schools. This prevents a possible gradual introduction, since teachers stress and drill usually the more probable examination themes, neglecting simply what is not important in this respect.

One may count on the natural curiosity of the pupils. If we do, we must find situations apt to motivate the children, and to put life into the theory. This seems to be a hard problem, since the allotted time is usually too short. Realistic examples are awfully complex, the simple ones are mainly artificial. We should find statistical populations, which can be easily sampled, inspected in a short time, and have time and space invariant stochastic character, providing a sound ground for the "average teacher". There are not so many such populations.

Field trips and case studies do not seem appropriate for our purposes, however desired they are. They can hardly be embedded into the framework of a unified centrally governed school system. We are furthermore convinced that only enthusiastic and knowledgeable teachers can treat them successfully.

Finally, one more observable problem: 16 year-old children can make statistical inference faster and more easily than the 18 year-old ones, if none of them had earlier statistical education. This fact contradicts the attempts to postpone the teaching of statistics to the last grade.

WHAT CAN WE DO NOW?

Under the present circumstances teaching of statistics should be accomplished within mathematics as part of the compulsory education with a possible facultative independent statistics course at the final grade of the secondary school. Topics in statistics and probability are to be dealt with from the very beginning.

For the time being we are forced to follow "The less the more principle", considering the present state of the teachers' knowledge. Treat just that part of statistics which harmonizes with traditional mathematics. A large part of descriptive statistics satisfies this condition. There are other problems, although usually neglected in introductory courses, which are common exercises in mathematics, but help to interpret statistical terms, e.g. the minimum-problems leading to the arithmetic mean and to the median in a statistical population are straightforward exercises in traditional mathematics, at the same time they clearly show one aspect of why they are considered as averages and why they are linked to the variance and to the standard deviation.

I consider it important to clarify the role of the pair mean-variance in representing statistical data, i.e. Tschebyseff-inequality for statistical data. No mathematics teacher objects to its teaching if properly handled (for a possible formulation, see Example 2 in the Appendix). This gives a bit of insight into the magic 3-sigma-rule, one of the most frequent cook-book recipes.

Teachers do not object to statistical inference problems if they can be solved on a common-sense basis, without sophisticated methods. One widely used example in this respect is breaking a simple substitutional cipher. This highlights an important difference between statistics and mathematics: statistical work may include some shrewd non-mathematical guesswork. A short cryptotext shows the difficulties with a small sample. A number of similar problems based on written languages can be considered equally well.

Many statisticians insist on dealing with tests of simple versus simple hypotheses. Some of them find that these must be preceded by providing the children with a sound background in probability. If we want to develop a proper attitude to statistical problems, this is not necessarily the case. There are simply formulated test problems which can be successfully attacked by "ad hoc" procedures proposed by the class, and the goodness of the decision rules can be experimentally investigated. Such a problem is stated in Example 3 in the Appendix.

The teaching of probability should concentrate on the notion of the random variable, in a possible strong connection to some statistical problem. At secondary level, it is difficult to design and perform a consistent and didactically acceptable course where probability is defined as a function on sets. Therefore we are against both axiomatic and set-theory approach, meeting the experiences in the Netherlands (see [2], p.53). The way, we propose, looks like this:

- Make a prediction of the distribution of a simple random variable
- Collect data
- Compare the observed data with the prediction at an intuitive level

This concept is illustrated by Example 4 in the Appendix.

WAYS TO SOLVE THE "TEACHER PROBLEM"

The 4th Round Table Conference on the Teaching of Statistics of ISI has worked out a recommendation in this respect. Let us recall a part of it ([2], p.18): "Teachers should be assembled in seminars for a few days - a week, say - and work together on a set of lessons already experimented to analyze, criticize, modify and adapt them... in those seminars, some of the time could be used to familiarize them with some real statistical work".

The loosely stated duration of such seminars becomes more concrete in a work of V. Barnett [1], who claims: "Courses of about 40 hours with plenty of practical work and student participation surely represent the minimum contact necessary to begin to bring teachers to a position where confidence and ability may develop".

It should be added that the best way to encourage them to teach the subject is to provide them with very detailed manuals, containing an almost minute-to-minute description of the lessons, and listing a wide variety of possible traps with solutions.

CONCLUDING REMARKS

Despite the great difficulties, elements of statistics can be included in the compulsory curricula of general educating secondary schools. For the time being we have to be self-restraining regarding the subject matter of instruction. The chances of its gradual extension are fair. It is vital, in this respect, to test carefully planned projects by a number of "average teachers" and not by enthusiasts. Teachers, participating in trials of the projects, should collect and formulate their experiences and problems. These, together with appropriate comments, and a detailed description of the project, should be widely circulated. An ambitious curricula, or a compulsory, independent statistics course may be introduced just after having re-trained a considerable number of the teachers.

References

[1] Barnett, V. (1980) *Statistical Education in Schools in England and Wales.* Technical Report for ISI TOTSAS, University of Sheffield.

[2] Breny, H. Ed. (1976) *The Teaching of Statistics in Schools*. Proceedings of the 4th International Statistical Institute Round Table Conference on the Teaching of Statistics.

[3] Nemetz, T. (1979a) *Pre-University Stochastics Teaching in Hungary*. Technical Report prepared for ISI TOTSAS.

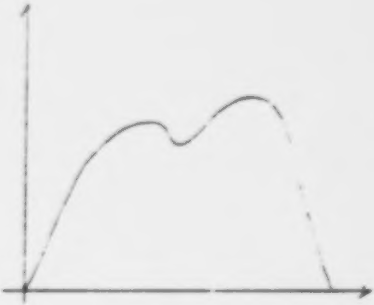
[4] Nemetz, T. (1979b) Stochastics in the Hungarian Schools: Pre-Univeristy level. In: *Convegno l'insegnamento pre-universitaro della statistica con particulare riferimento alla scuola secondaria superiore*. Bressanone, 19-21 settembre 1979.

APPENDIX

Example 1: A prospective examination-test question in biology.

The figure shows the distribution of a feature. Choose the most realistic statement from the following one!

A) There is no feature with such a distribution
B) The figure shows the typical distribution of a quantitative feature
C) It shows the typical distribution of a qualitative feature
D) It shows the distribution of a quantitative feature based on a very small sample
E) It shows the distribution of a qualitative feature based on a very small sample



Example 2: The proof of Tschebyseff inequality does not frighten teachers of mathematics.

Claim: Let $x_1, x_2, \ldots, x_n$ be a sample of non-negative numbers, $\bar{x}$ be the arithmetic mean, and $A > \bar{x}$ an arbitrary number. Then there are at most $\frac{n\bar{x}}{A}$ sample elements which are not smaller than A.

Proof: $n \cdot \bar{x} = x_1 + x_2 + \ldots + x_n =$
$=$ (the sum of x_i's which are less than A) $+$
$+$ (the sum of x_i's which are not smaller than A)
$\geq$ (the sum of x_i's which are not smaller than A)
$\geq A \cdot$ (the number of x_i's which are not smaller than A), etc.

Interpretation: The proportion of elements in a sample of non-negative numbers which do not exceed a given number A, where $A >$ sample mean $\bar{x}$, is at least $1 - \frac{\bar{x}}{A}$.

Now, if we have a sample $y_1, y_2, \ldots, y_n$ containing arbitrary real numbers, and the arithmetic mean $\bar{y}$ and the variance S^2 are known, then we can apply our result to the sample $(y_1-\bar{y})^2, (y_2-\bar{y})^2, \ldots, (y_n-\bar{y})^2$, $A = \lambda^2 S^2$.
Since the arithmetic mean of this new population is S^2, we get:
The proportion of the elements y_i in the sample $y_1, y_2, \ldots, y_n$, for which $(y_i-\bar{y})^2 \leq \lambda^2 S^2$, is at least $1 - \frac{S^2}{\lambda^2 S^2} = 1 - \frac{1}{\lambda^2}$.

Interpretation:

If we know only the mean $\bar{y}$ and variance S^2 of a sample $y, y_2, \ldots y_n$ of reals, we can be sure that at least $(1 - \frac{1}{\lambda^2})n$ elements belong to the interval $[\bar{y} - \lambda S, \bar{y} + \lambda S]$
Here, one can take concrete values for λ, e.g. $\lambda = 3$.

Example 3: A decision problem.

A statistical population contains pairs of type A and type B objects. Both types of objects have 4 characteristics. These characteristics can be measured in a given order.

A sample of 10 elements was taken and is given here. The first column gives the label assigned to the given pair. Another person did not pay attention if he was measuring a type A or a type B object. In his records, therefore, the ordered quadruples may have got interchanged. Your task is to guess if it happened or not. Write AB into the last column if you think that they are not interchanged, and write BA otherwise!

No.	TYPE A				TYPE B			
7	11	7	1	1	15	2	5	1
8	14	1	11	1	21	3	27	9
9	29	6	1	1	9	20	18	3
10	19	1	1	1	14	5	6	48
11	2	12	1	1	1	1	1	1
12	14	1	1	1	30	1	1	40
13	2	20	1	7	1	1	1	64
14	8	10	1	1	26	7	2	27
15	6	72	10	1	19	11	5	3
16	3	1	1	8	16	1	8	17

No.	1. Data Set				2. Data Set				?
17	6	26	32	17	32	24	1	3	
18	12	1	1	2	19	4	4	7	
19	1	18	9	16	19	1	1	3	
20	2	35	5	5	1	1	1	31	
21	9	12	3	29	21	98	6	1	
22	8	33	22	47	1	1	1	3	
23	8	9	10	1	5	1	3	1	
24	8	26	11	6	34	1	15	3	
25	18	4	25	7	30	1	1	16	
26	6	12	17	3	15	7	2	10	
27	7	23	4	3	4	3	5	3	

For the reader's information: type A data was obtained in the following way. We have taken the book [2], and opened it at the page given in the first column. Then we have counted the letters and spaces on that page till the first letter *T*, to get the first "measurement". Starting from here we have counted again till the first *H* to get the second one, then till the next *E*, finally, till the next SPACE. The key-word STAR was used to get type B data. Finally, the interchanges were made on a coin-tossing basis.

Example 4: The problem of two coins.

Concerning the number of heads when tossing two coins there are usually two models suggested in the classroom:

$$A = \begin{Bmatrix} 0, & 1, & 2 \\ 1/3, & 1/3, & 1/3 \end{Bmatrix}; \qquad B = \begin{Bmatrix} 0, & 1, & 2 \\ 1/4, & 1/2, & 1/4 \end{Bmatrix}$$

The question, which is the "better", is dealt with via experiments. It is not boring to collect some 100-120 observations. A typical result;

$$E = \begin{Bmatrix} 0, & 1, & 2 \\ \frac{28}{120}, & \frac{58}{120}, & \frac{34}{120}, \end{Bmatrix}$$

I was surprised to learn, and this is why this example was given, that children of 16 years of age frequently come up with the following decision-rule:

The probabilities in models A and B, the relative frequencies in E can be viewed as points in the 3-dimensional space. We will say A is better than B, if E is nearer to A than to B, and vice versa.

STATISTICS FOR THE 16-18 YEAR OLDS: A DISCUSSION

A Hanley, Northampton

There seems to be agreement amongst statistical educators that appropriate statistics courses should be provided for all 16-18 year old students. Professors Eicker and Nemetz have described what form this statistics curriculum could take. They have also given reasons to justify the introduction of statistical education which could be presented to the education authorities who, in many countries, decide centrally what the components of the school curriculum shall be. A major task is then to persuade these authorities that statistics merits a place in the curriculum even if the time allocated to other studies has to be adjusted to accommodate it. When this task of persuasion is accomplished schools will be obliged to introduce statistics into their curricula.

In the United Kingdom the task has a different focus. There is no centrally prescribed curriculum so that schools are free to choose and plan their own. Adjacent schools could be providing different curricula, each chosen by the staffs of the schools to realise their personal views on what constitutes the best course for students. In practice there is not the variety there might be owing to the influence exerted by the public examinations. These examinations are so highly regarded by parents, employers, teachers and hence by students that courses not leading to a public examination are regarded as of very low status and value. Schools therefore tend to offer packages of courses based on public examination syllabuses which leads to a fair degree of uniformity in the range of subjects on offer whilst variety of course content arises from the selection of a particular subject syllabus.

Changes in a school's mathematics curriculum usually follow an initiative by the Head of the Mathematics Department who may be influenced by a school colleague, by a local authority adviser, by an inservice course, by a professional association or by a national "fashion". Suggestions about desirable curricular changes are regularly made in reports published by the Government following the studies of Committees established to examine particular aspects of the educational system or following investigations carried out by Her Majesty's Inspectors of Schools. Such suggestions are frequently not acted upon. For example, the Crowther Report [1], published in 1959, had this to say about "numeracy", a word invented by the Crowther Committee and used often at this Conference.

> Literacy has long been important, and its value is as great as ever. Just as by "literacy", in this context, we mean much more than its dictionary sense of the ability to read and write, so by "numeracy" we mean more than mere ability to manipulate the rule of three. When we say that a scientist is "illiterate", we mean that he is not well enough read to be able to communicate effectively with those who have a literary education. When we say that a historian or a linguist is "innumerate" we mean that he cannot even begin to understand what scientists and mathematicians are talking about. The aim of a good Sixth Form* should be to send out into the world men and women who are both literate and numerate. It is perhaps possible to distinguish two different aspects of numeracy that should cocnern the Sixth Former. On the one hand is an understanding of the scientific approach to the study of phenomena - observation, hypothesis, experiment, verification. On the other hand, there is the need in the modern world to think quantitatively, to realise how far our problems are problems of degree even when they appear as problems of kind. Statistical ignorance and statistical fallacies are quite as widespread and quite as dangerous as the logical fallacies which come under the heading of illiteracy. The man who is innumerate is cut off from understanding some of the relatively new ways in which the human mind is now most busily at work ..... However able a boy may be, he cannot reach a Sixth Form level of numeracy except on the foundation of a Fifth Form* level; but, if his numeracy has stopped short at the usual fifth form level, he is in danger of relapsing into innumeracy. It is now one of the most important tasks of the Sixth Form to ensure that no boy or girl leaves school as innumerate as most have done in the past, and as far too many do even today.

Despite this exhortation, few schools over 20 years later have introduced the types of Sixth Form course envisaged by Crowther. Indeed, a prime concern of this Conference is that such courses should be developed for the 16-18 year olds.

---

* The Sixth Form consists of the 16 to 18 year old students in secondary education. The Fifth Form is the group of 15 to 16 year old students.

One of the reasons why advocated changes do not take place is that central government does not usually designate any portion of the grants to local authorities specifically to curriculum development. The latest report [2] has, inside the front cover, the following "disclaimer" (reminiscent of the Government Health Warning appearing on packets of cigarettes sold in the UK) "No assumption can be made about government commitment to the provision of additional resources as a result of this report." Allocation of financial resources to implement curriculum development requiring them is therefore a matter of local decision.

Whether innovation is directed centrally or chosen by the individual school the support that will be demanded from those already involved in statistical education will be substantially the same. Most teachers are not prepared by initial training to teach statistics and even mathematics specialists may have little statistical background. Inservice courses could be provided systematically to train teachers over a suitable period of time. (In the UK there is no commitment to compulsory inservice training that would ensure teachers were exposed to this). Suitable teaching materials are scarce but they are slowly being produced. Here there are the Schools Council's Project on Statistical Education booklets, a few Open University courses, some BBC schools' broadcasts and associated materials, and soon there will be materials coming from the Leverhulme Statistical Education Project 16-19. There is a taxing exercise in front of all of us to produce all that is going to be demanded to introduce sound statistical education into the curriculum of 16-18 year olds.

A recent survey of mathematics teaching in schools in England and Wales [3], investigated the mathematical needs of adults in their work and in everyday life. It was found that many adults were very apprehensive about being asked to face simple mathematical tasks - indeed, a good proportion of those approached refused to co-operate when they knew the topic was mathematics. Many of those questioned could not cope with percentages or simple averages nor read a timetable correctly. They ascribed their lack of confidence and understanding to a variety of reasons but a major one was the irrelevance to them of the work they had been asked to do at school. Is there a danger that courses in statistics might be subject to the same criticism? If so, what should we do to avoid this pitfall?

Lastly, we are a gathering of enthusiasts. Many of the teachers who may be required to tackle new courses will have no better, perhaps worse, statistical backgrounds than the students they teach. How can they be given the practical experiences of relevant statistics themselves so that they may teach their students with the conviction and insight based upon the confidence of personal experience? A number of projects have tried to produce resource materials which are "teacher-proofed" in the hope that teachers cannot misuse them. Can we adopt the positive attitude of designing materials that will be "teacher-improving"?

References

[1] HMSO (1959) 15 to 18: A report of the Central Advisory Council for Education (England)

[2] HMSO (1982) Education 5 to 9: an illustrative survey of 80 first schools in England.

[3] HMSO (1982) Mathematics Counts: Report of the Committee of Inquiry into the Teaching of Mathematics in Schools under the Chairmanship of Dr. W.H. Cockcroft.

TOPIC 4: TEACHING STATISTICS TO NON-STATISTICIANS

Organiser: Dr W G Gilchrist

WE MUST PREACH WHAT IS PRACTISED

A S C Ehrenberg, London Business School

SUMMARY

Much of what we teach in statistics is seldom used in practice.
Much of what students need, we do not teach.

In teaching non-specialists, we need to cover the procedures they will come across in their other studies and subsequent work. Since many courses and introductory texts do not do so, students are often disillusioned with them and regard such courses as irrelevant.

I now describe some topics which we often teach unnecessarily (or with the wrong emphasis) and some which we mostly do not teach at all even though they are often needed. They range from deep matters like statistical inference or causation to more hum-drum (but important) ones like means, medians and modes.

FREQUENCY DISTRIBUTIONS OR TABLES

We spend much time teaching frequency distributions, both in dealing with observed data (grouped and ungrouped readings, histograms, etc.) and in telling students about theoretical models like the Normal, Poisson, or Binomial distributions. However, students seldom come across frequency distributions in their readings or practical work. This also applies to teachers - in my experience, when asked when they last used or saw a frequency distribution, teachers generally refer to some teaching occasion rather than to anything practical.

What students see instead is aggregated data: averages, totals, percentages and so on. In particular, they do so in the form of structured tables, where each entry is some such summary statistic. However, we hardly ever teach students how to analyse and interpret tables.

I am not suggesting that frequency distributions should not be taught at all. They are usually the basis of the aggregated statistics people do see. They also occur in dealing with deviations from a model, and they arise in random sampling (at least hypothetically). But if we want to avoid students being puzzled at what they are being taught, we ought to explain to them the real role of frequency distributions: why they do not often see them explicitly and why they nonetheless ought to know something about them.

MEANS, MEDIANS AND MODES

We spend a good deal of time teaching the calculation of means, medians and modes. Yet the latter two measures are virtually never mentioned again in the rest of the course or textbook, nor do they occur much in students' other reading.

Medians and modes are helpful concepts in understanding observed data. But they are seldom used in analysis, e.g. in comparing different sets of data or in model-building. This is not a value judgement, but a matter of fact. It ought to be reflected in our teaching.

THE MEAN DEVIATION AND THE STANDARD DEVIATION

Many introductory texts do not mention the mean deviation at all. Others denigrate it. Yet faced with a range of numbers like the deviations from a fitted model - e.g. 3, -1, 4, 0, 1, -3, 3, 0, etc. - we generally note that they are roughly 2. We form a mean deviation either in our head, or more precisely on paper, $15/8 \doteqdot 2$.

The standard deviation has well-known advantages, being both easier to compute from raw data, especially for large n, and also vastly more amenable to theoretical developments. But for the beginner the standard deviation only stops being a rather mysterious abstraction if we lead into it as a development from the mean deviation.

Both as a teaching aid and as something which most people often use in practice, the mean deviation is sadly neglected.

RANDOM SAMPLING VERSUS OBSERVATIONAL DATA

Random sampling in general and randomised experiments in particular are often treated in introductory texts as if they were the only ways of obtaining proper data.

But randomised experiments are only just over 50 years old. In practice, observational studies predominate, with all their additional problems. Our teaching should reflect this.

Further the vast majority of data are not collected by any method of explicit random selection. Each set of data is in effect a small ad hoc population. Yet such data are not necessarily bad. Measuring the whole population may be wasteful, but it is not wrong.

What is needed is the notion of generalisation. But this is hardly ever discussed. I have not seen the word "generalisation" or anything like it in the index of a statistical textbook, let alone in the body of the text. Yet the idea of a result which is known to hold under a wide range of different conditions is basic to ordinary science and technology.

BEST-FIT EQUATIONS VERSUS LAWLIKE RELATIONSHIPS

Virtually all our teaching of the relationships between two or more variables is in terms of finding the best-fitting equation for a single set of data. But this is not what students do (or should) come across in their ordinary studies - a result which was only established once and which has never been repeated.

In contrast, fitting an equation to two or more different populations (or samples therefrom) is not discussed in our literature. The derivation and interpretation of the lawlike relationships of science does not seem to be regarded as a fit occupation for a statistician.

PROBABILITY VERSUS CAUSATION

Probability figures large in many statistics courses and texts. Some even start by saying that statistics is about probabilistic data.

Yet having been taught the laws of probability (usually in terms of games of chance with equally likely outcomes - a highly improbable concept that never occurs elsewhere in life or science), students find that neither they nor their introductory text make any further use of probabilities. (The basic notion of random sampling does not require a detailed mastery of the probability calculus).

In contrast, the process of trying to move towards causal interpretations of observed data is seldom discussed in the usual texts.

EXPLORATORY ANALYSIS VERSUS REPETITIVE STATISTICS

Most situations treated in traditional statistics are exploratory. This was so even prior to John Tukey's emphasis on EDA.

Each new set of data is to be analysed as if there were no prior knowledge of other similar data. This for example underlies our usual notions of inference ("Is this correlation significant?") and that of fitting a regression line ("How does y vary with x in this data?").

But in practice, most of the data professional people come across are repetitive. They have seen similar data before, usually week-in week-out. They therefore have much prior knowledge and understanding of the phenomena (except perhaps in the case of statistical consultants).

DISCUSSION

In this paper I have briefly outlined a number of points where what is taught (i.e. what is in our textbooks) differs greatly from what students actually need.

I believe that we must change what we teach, to cover the kinds of statistics that students come across in their main subject-areas and their practical work. This seems a more important topic than whether courses should be taught by statisticians or subject-matter specialists (to which there can hardly be a single general answer).

The re-orientation I am suggesting will also overcome statistical teachers' continuing lack of practical teaching materials. They often report that they cannot extract practical illustrations or case-histories of using "Technique X" from statistical practitioners. But as Professor Gerald Goodhardt has pointed out, the reason is simply that although Technique X is in our textbooks and is taught in our courses, it is practically never used. That is why there are no practical examples.

In adapting our teaching, some temporary compromises will be needed. There are two reasons. First, individual teachers will not always find it easy to give up instantly much of what they have been teaching for many years. Secondly, we may have to go on covering items that are required by syllabuses or examination systems (which often take longer to change). Nonetheless, we can work at making our texts and teaching materials more relevant.

TEACHING STATISTICS TO NON-STATISTICIANS - AN APPLICATION ORIENTED APPROACH

G Sadasivan, Indian Agricultural Statistics Research Institute

INTRODUCTION

Teaching statistics to non-statisticians is more difficult than teaching the subject to statisticians. Some methods of teaching statistics to statisticians are given in Kempthorne (1980), Mosteller (1980), Noether (1980) and Tukey (1980). The role of statistics in the teaching of agricultural scientists is given in Sadasivan (1966). In the author's opinion the classroom type of teaching is much better for such personnel than the seminar type of teaching. This is specially so if the performance of the trainees is to be assessed by any type of test subsequently. It may be noted that the statistical topics to be taught to different categories of non-statisticians are different. The methods of teaching these topics should be oriented towards the methods of assessment of the trainees. Teaching with detailed mathematical derivation would be troublesome to the non-statisticians. For instance, for introducing statistical quality control in textile industry the mill owners are to be instructed in very simple terms how statistics can be used to benefit production of textiles of acceptable quality level. This is a seminar type of teaching non-statisticians. This type of teaching is of relevance in Government Departments as well, where the boss does not know anything in conventional statistics, but has to take decisions using statistical data in his day to day activities. The agricultural scientists in India are also a group of non-statisticians who know next to nothing in statistics and it is the job of the statisticians to explain, say, the benefit of using a complicated experimental design for such an audience. For instance, the present generation of agricultural scientists in India did not want to use anything more than randomised blocks or split-plots for their field experiments. In many cases they used to lay out their experiments without using any statistical considerations and later on approached a statistician to squeeze out as much information as possible from their experiments. Although this gives rise to new problems in statistics, it is not a happy situation. One way to prevent such mishaps is to impart appropriate statistical training to such groups of non-statisticians. But a word of caution is necessary here. The training of such personnel should not be highly mathematically oriented lest they totally dispense with the services of statisticians.

It is clear from the exposition above that there are different categories of non-statisticians as (1) ordinary scientists (2) agricultural scientists (3) engineers and technologists (4) business scientists (5) management scientists (6) administrators (7) farmers (8) other laymen and (9) the general public. The teaching method need not be the same for each of these categories because the educational level especially in mathematics is different among these categories. Hence the teaching method has to be oriented towards the needs of the particular subgroup we are considering. We have also to consider whether there can be a general method of teaching applicable to all non-statisticians.

Some of the organisations involved in teaching statistics to non-statisticians in India are the Indian Statistical Institute, Indian Agricultural Statistics Research Institute, Agricultural Universities, Indian Institutes of Technology and regular Universities. Teaching non-statisticians in these organisations calls for inter-disciplinary collaboration. As far as collaboration between agricultural scientists and statisticians is concerned, the scientists will always have an edge over the statistician.

Teaching experimental results to farmers also falls in the class of teaching statistics to non-statisticians. Since all experimental results are expressed statistically, usually only the interpretation of statistical analysis of the results is given to the farmers. They are taught with the help of diagrams, charts, tables, etc. and with the help of projectors by the extension specialists. The data are derived from comprehensive experiments in the cultivator fields or from the results of the experiments in research stations. The teaching involves results of precise and careful data analysis. If the literate farmers are given the rudiments of the concepts in statistics, they will be able to appreciate the experimental results much better. To make this programme effective the agricultural extension workers are to be trained adequately in statistics.

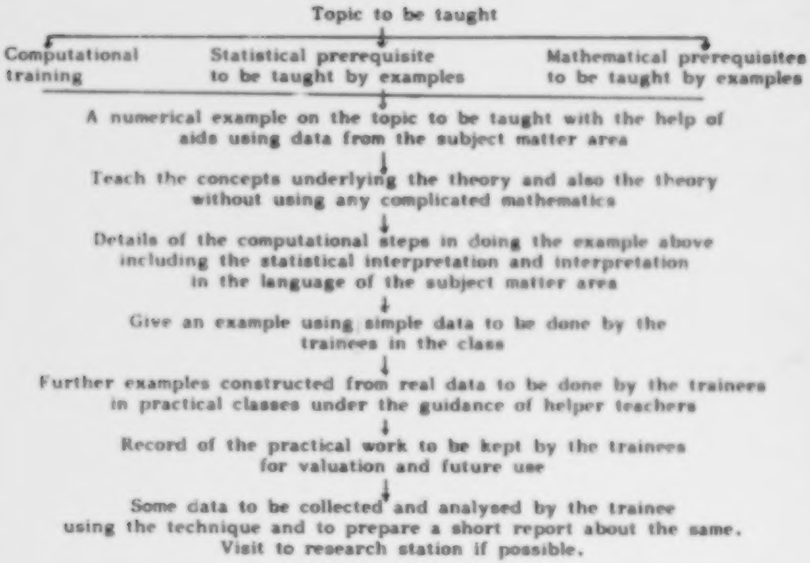
There are many statistical training courses for non-statisticians at the Indian Agricultural Statistics Research Institute, New Delhi. In 1958 we framed two courses for the benefit of agricultural scientists. These were later reorganised into six courses for experimental statistics. Then the necessary pre-requisite courses in Mathematics were framed especially for students in Agricultural Physics, Agricultural Engineering, Genetics, Agricultural Economics etc. It may be noted that the courses formulated for each of these disciplines contain the appropriate topic in statistics useful to them. The courses for Genetics students include Biometrical Genetics, Population Genetics and Biometrical Methods, courses for Entomologists include Tagging Models, Probit Analysis, Bioassay Techniques, courses for Mycologists include Contageous, Negative Binomial, Log-normal and other types of distributions apart from Elementary Statistical Methods. Similarly the courses for Agronomists include Advanced Design of Experiments, courses for Extension include Nonparametrics and Psychometric methods, courses for Economics include Econometrics. Other disciplines in Agriculture use elementary statistical methods. Similarly there are certificate courses in the Institute for training nominees from different states in India and abroad who had no formal training in statistics to enable them to perform their day to day statistical duties. These trainees belong to a wide spectrum, most of them matriculates or degree holders in Agriculture or Animal Sciences but by virtue of their jobs are doing statistical work in their parent departments. Some of the trainees have deficiencies in languages. There is arrangement to train them in English before their initiation to statistics courses. Apart from these there are short term courses in the Institute which impart training in statistics to Directors of Agricultural Institutes, FAO Trainees, personnel of the All India Services such as Indian Administrative Service, Indian Forest Service, Agricultural Research Service, etc. The method used by the author for training these personnel along with some of the other possible methods are detailed in this article. Apart from these, the author is associated with the Food Scientists in India who use a number of non-parametric methods in their training programme in statistics. It may be noted here that the training imparted to Indian Administrative Service and other probationers is of a general nature which helps them to understand the working of the different sections and organisations dealing with statistics under the Indian Government. Further I.A.S.R.I. had a programme to train non-statisticians to professional statisticians. Under this programme a graduate in Agriculture can be trained for the M.Sc. degree in Agricultural Statistics. The first year of the course is used for teaching supplementary mathematics. The second and third year are used for training in statistics including a research project.

Now a word about teaching non-statisticians in Universities. The methods of teaching used below can also be used in this case. Usually, for teaching in Universities real data are not used. The students are not given any project work.

Advisory work to Agricultural Scientists for conducting their research is a type of training of the non-statisticians in the art of data collection. This involves training in planning and analysis of their field and laboratory experiments in such a way that data on different characters can be collected and analysed. In these cases knowledge of programming is an advantage to the non-statisticians.

THE MAIN TEACHING METHOD

Here we discuss an application oriented method of teaching statistics to non-statisticians. It is described by chart as follows:-



Topic to be taught

Computational training

Statistical prerequisite to be taught by examples

Mathematical prerequisites to be taught by examples

A numerical example on the topic to be taught with the help of aids using data from the subject matter area

Teach the concepts underlying the theory and also the theory without using any complicated mathematics

Details of the computational steps in doing the example above including the statistical interpretation and interpretation in the language of the subject matter area

Give an example using simple data to be done by the trainees in the class

Further examples constructed from real data to be done by the trainees in practical classes under the guidance of helper teachers

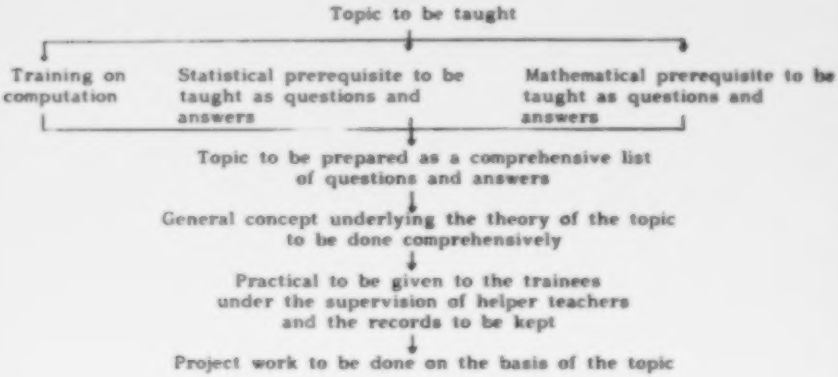
Record of the practical work to be kept by the trainees for valuation and future use

Some data to be collected and analysed by the trainee using the technique and to prepare a short report about the same. Visit to research station if possible.

The method can as well be described as follows: First decide about the topic to be taught. Then teach the required pre-requisite Mathematics and Statistics as supporting courses in a simple way by doing simple exercises. Then construct a numerical example from the subject matter area for the topic we decide to teach and interpret the statistical procedure using the numerical example. If charts or other teaching aids are helpful they can be used for this teaching. Proceeding from this we can state the general statistical concepts underlying the procedure along with the necessary formulae, stating clearly what the symbols in the formulae stand for. Then the computational steps involved in doing the practicals relating to the procedure are to be given in such a way that the results can be calculated easily by the substitution of the unknowns in the formulae from the numerical. A statistical interpretation of the results obtained should be given next. This interpretation is further translated in the language of the subject matter area for the convenience of the non-statisticians. The computations in the above should be such that they can easily be performed with the hand or with the help of a facit calculator. In cases where the non-statistician knows programming the computations are made using an electronic computer. But this procedure will help the scientist to comprehend fully the implications underlying the statistical method only if he knows programming. Now give a small example to the trainee as classwork. Then good practicals based on the procedure are to be constructed from the subject matter area and given to the trainee. These computations are demonstrated by helper teachers. This will help the non-statistician to keep a good record of the work done by him for future use in all his scientific endeavour. If possible, some data should be collected, analysed by the trainee from field or laboratory or any other source using the technique under study and a short report written. Visits to research stations practising the methodology will further help the non-statistician to understand the method.

OTHER TEACHING METHODS

The teaching methods should be oriented towards the method of assessment. The method given above would be useful if the questions are comprehensive in nature. But the recent trend in assessment method is to ask objective type questions. For this method of assessment a teaching method designated Method II is diagrammatically as follows:-


Topic to be taught
Training on computation
Statistical prerequisite to be taught as questions and answers
Mathematical prerequisite to be taught as questions and answers
Topic to be prepared as a comprehensive list of questions and answers
General concept underlying the theory of the topic to be done comprehensively
Practical to be given to the trainees under the supervision of helper teachers and the records to be kept
Project work to be done on the basis of the topic

In words Method II can be described as follows. First the prerequisites to the topic can be taught in the form of questions and answers. The next step is to prepare a comprehensive list of questions, the answers to which would cover the topic to be taught. Next the general concepts underlying the topic can be taught followed by a list of questions for practical work supervised by helper teachers. The rest is as in Method I. In this method of teaching much expertise is required for proper drafting of questions and their possible answers. The questions can be oriented towards practical things as well as theory. In teaching non-statisticians, the quantum of practical questions should be more than that for theory.

In the following Method III yet another technique of teaching non-statisticians is discussed. Here the topic to be taught is assigned to a trainee. Under the supervision of the statistician the student is to prepare the teaching material and teach the topic to the trainees. The statistician should be present in the class to resolve any difficulties while teaching. The order and technique of presentation can be either as in Method I or Method II. It may be noted that Method III facilitates better discussion.

REQUIREMENTS TO FACILITATE BETTER TEACHING

In the following we discuss some requirements other than those given earlier for better teaching. The courses for training of non-statisticians should be introduced from elementary levels to advanced levels. For training such personnel teaching aids are very helpful. They include, graphs, diagrams, projectors, calculators, television programmes, computerised courses, practical classwork, class notes, statistical and mathematical tables, boards, chalk, library and other equipment. Each category of non-statistician requires a special area in statistics. His training should be oriented towards this need. This specialised training is very useful for developmental research in general. There are special statistical publications which are of interest to non-statisticians. These must be listed and brought to their notice. Non-statisticians should be encouraged to collaborate in statistical research projects applicable in their areas. This will ensure better applicability for the methods evolved from such projects. Non-statisticians can participate in statistical research seminars which are of relevance to their specialised areas. Teaching of non-statisticians requires inter-disciplinary collaboration. If non-statisticians develop a supercilious attitude such collaboration is difficult to achieve. In training non-statisticians there should be specific reasons for including statistics as a part of the curriculum. Every scientist should be taught data gathering, principles of design and survey planning. In teaching non-statisticians case studies can be utilised. The needs of course contents for different types of non-statisticians are different. Some of these were briefly discussed in the introduction.

EXAMPLES OF TEACHING METHODS

Example I: Let us consider the example of teaching the binomial distribution and some of its applications to agriculture.

Step I - teach the following prerequisite mathematics and statistics to the trainees: proportions and percentage; square roots; theory of indices; frequency distributions - discrete and continuous; central tendency and dispersion, calculation of mean and variance from grouped data; probability - addition and multiplication rules; mathematical model building; sampling for proportion from finite and infinite populations; estimates and their properties. Train the participants in computational techniques as well. Then teach a numerical example like the following:-

Step 2 - in an experiment on wheat we desire to find the proportion of wheat plants affected by wheat rust along with its error. Divide the area into micro-plots of the same shape and size, each plot containing $n = 10$ plants. Select $N = 100$ such plots at random from the tract which has in all $N^1 = 1,000$ such plots. The distribution of the number of affected plants per plot is as follows:

Number of affected plants per plot (x_i)	0	1	2	3	4	5	6	7	8	9	10	Total
Number of plots (f_i)	5	7	12	15	15	20	15	6	3	1	1	100

Assume that the probability p, for incidence of wheat rust is constant over the area. The problem is to estimate this probability along with the appropriate error assuming that the binomial distribution gives a close fit to the data. First assume that the $N = 100$ sub-plots come from an infinite population. If the binomial gives a close fit to the data, the mean of the data $\bar{x} = np = 10p$, the mean of the binomial. Thus an estimate of p is $\hat{p} = \bar{x}/10 = 4.11/10 = .411$. A biased estimate of the error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/10} = .156$. An unbiased estimate of the error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/9} = .164$. Since the above data gives the exact distribution of $\hat{p}$, the exact error of $\hat{p}$ works out directly from data as

$$\sqrt{\frac{1}{99}\left[\sum_{i=1}^{n} p_i^2 f_i - \frac{(\Sigma p_i f_i)^2}{100}\right]} = .212 \text{ where } p = x_i/n.$$

These values are different since they have been obtained by different theories. Again, considering that the sample is coming from a finite population of $N^1 = 1,000$ plots, using Cochran (1953, p.34) an unbiased estimate of the variance of $\hat{p}$ is

$$\frac{(N^1 - N)}{N^1 N(N-1)} \left[\Sigma p_i^2 f_i - \frac{(\Sigma p_i f_i)^2}{N}\right] = 0.000407$$

giving an estimated standard error of .0202. If we use the concept that the sample we have is a simple random sample of 1,000 plants from 10,000 plants,

$$v(\hat{p}) = \frac{(N^1 - N)}{(nN-1)N^1}\, \hat{p}(1-\hat{p}) = 0.000218$$

and hence estimated standard error of $\hat{p}$ = .0148. This smaller value of the error should be expected since the procedure assumed here is quite different from the one used.

Step 3 now we interpret the theory behind the numerical example as follows. We have assumed that the binomial distribution gives a good description of the data. Thus there is a constant probability $\hat{p}$ of a plant getting affected by rust in the tract. Hence assuming sampling from infinite population, the expected frequencies of samples with 0,1,2,...10 affected are given by the successive terms in the expansion of 100 $(\hat{q} + \hat{p})^{10}$ where $\hat{q} = 1-\hat{p}$. The mean of this distribution is 10 $\hat{p}$. The estimated standard error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/10}$. An unbiased estimator of standard error of $\hat{p}$ is $\sqrt{\hat{p}(1-\hat{p})/9}$. From sampling theory for finite populations, since every sampling unit contains the same number n of plants an estimate of the variance of $\hat{p}$ is

$$\theta(\hat{p}) = \frac{N^2-N}{N^2 N} \cdot \frac{1}{(N-1)} \sum_{i=1}^{N} (p_i-\hat{p})^2 f_i .$$

If the plants are taken as the ultimate units and if we are drawing n N elements out of nN^2 elements at random, another estimate of variance is given by

$$\theta(\hat{p}) = \frac{(N^2-N)}{(nN-1)N^2} \hat{p}(1-\hat{p}).$$

Step 4 - The different formulae for computation can easily be derived from the above. Using these in the numerical example, the mean proportion of incidence of the disease in the tract is .411 with the most appropriate estimate of error .202. Hence the recommendation to the extension scientists is as follows. At a particular time of crop growth under some suitable conditions rust disease can occur in the wheat crop. Knowing the extent of incidence appropriate spraying can be made to save the plants from disease. The recommendation to farmers can be as follows. Find the proportion of plants in the tract affected by the disease. If it is above danger mark as given by the mycologist resort to spraying. Thus the interpretation of the results goes outside the ambit of statistics.

Other steps in teaching Method I can easily be followed. Method II can also be used for teaching this example.

Example 2: The data used in this example are adopted from Davidson and Beaver (1977) in preference testing for packed food mix labelled 1,2,3,4. they are given below.

Pair (i,j)	Number of repetitions (n_{ij})	Number of preferences for i $a_{i \cdot ij}$	Number of preferences for j $a_{j \cdot ij}$
-	-	-	-
(12)	85	37	48
(13)	85	47	38
(13)	83	52	31
(14)	84	49	35
(24)	82	48	34
(34)	85	45	40

The problem is to teach the non-statistician the generalisation of the Bradley - Terry Model by Dykstra (1960) using the above data.

First teach the computational methods and also the following prerequisites in a simple way.

Mathematics: Matrix algebra; inversion of matrices; logarithms, solutions of linear equations; differentiation; maxima and minima, Lagrangean multipliers.

Statistics: probability and likelihood; maximum likelihood; concept of parameter-free procedures; hypothesis testing; chi-square distribution; efficiency concept in non-parametrics.

These topics can be taught by providing prerequisite courses or independently.

Step 1 - We teach by Method I. The problem here is to obtain the ratings for the four packed food mixes using Dykstra's model. Under the model the estimates p_i (i=1,2,3,4) are obtained from the normal equations set up by the method of maximum likelihood by an iterative procedure. We programmed the iteration and the solutions converged to p_1 = .26398837; p_2 = .32149677; p_3 = .21121396; p_4 = .20330090.

Since the number of repetitions are large in the experiment an asymptotic test can be used to test whether p_i's differ significantly among themselves. The test statistics $- 2\ln\lambda$ in this case has been computed as 11.1028 with three degrees of freedom and is hence significant showing that the food mixes differ significantly among themselves. The statistic to test the appropriateness of the model to the data above has been computed as 1.0519 which is distributed as a chi-square with 3 degrees of freedom and is hence not significant showing that the model is quite appropriate. The 3 x 3 asymptotic variance-covariance matrix for testing any contrast among p_i's is:

$$\begin{bmatrix} .00107 & .00028 & .00054 \\ .00028 & .00109 & .00026 \\ .00054 & .00026 & .00133 \end{bmatrix}$$

Step 2 - Now we discuss the theory behind this method. Under Dykstra's model it is assumed that t treatments are tried in t(t-1)/2 pairs, each pair being repeated an unequal number of times. It is postulated that corresponding to the t treatments T_1, T_2, ... T_t there exist parameters π_i for T_i (i = 1,2,....,t) such that $\pi_i > o$ and $\sum_{i=1}^{t} \pi_i = 1$. The parameters are further defined with the probability statement that, if X_i generally denotes an observation on a sample of T_i, $P(X_i > X_j) = \pi_i/(\pi_i + \pi_j)$ in the comparison of T_i with T_j, $(i \neq j)$. Under these formulations, the general likelihood under the model becomes

$$L(\pi_i) = \prod_i \pi_i^{a_i} \prod_{i<j} (\pi_i + \pi_j)^{-n_{ij}} \quad ..... (5.1)$$

where $a_i = \sum_{j \neq i} a_{ij}$ and $a_{ij} = 2n_{ij} - \sum_{k=1}^{n_{ij}} v_{ijk}$

$a_{ij} + a_{ji} = n_{ij}$ where $v_{ijk} = 1$ if $x_i > x_j$ and $v_{ijk} = 2$ if $x_j > x_i$ in the kth repetition of the pair (ij). Next we test the null hypothesis $H_o : \pi_i = \frac{1}{t}, (i=1,2,\ldots,t)$ against the alternative hypothesis H_1: no π_i is assumed equal to any π_j $(i \neq j)$. Then differentiating the natural logarithm of (5.1) with respect to π_i using appropriate Lagrange multipliers and denoting the maximum likelihood estimates by p_i we obtain $(a_i/p_i) - \sum_{j \neq i} [n_{ij}/(p_i+p_j)] = 0$ $(i,j = 1,2,\ldots,t)$ with the usual restriction $\sum_i p_i = 1$. For estimating p_i the above equation can be put in the form:

$$p_i = a_i/[\sum_{j \neq i} n_{ij}/(p_i+p_j)] \ldots\ldots\ldots\ldots (5.2)$$

Reasonable first estimates may be obtained by assuming that the p_j's are not too different from each other so that we may assume $p_j = p'$ for all $j \neq i$. Then

$p' = (1-p_i)/(t-1)$ and we obtain

$$p_i = a_i/[(t-1)\sum_{j \neq i} n_{ij} - (t-2)a_i] \ldots\ldots\ldots\ldots\ldots\ldots (5.3)$$

(5.3) can be used to get first estimates of p_i. These estimates can be substituted into the right-hand side of (5-2) to obtain second estimates. The procedure is iterated till two successive estimates are equal. The number of iterations required under this procedure is small provided the p_i's are not widely diverging. For testing H_o against H_1 the likelihood ratio depends on the statistic,

$B_1 = \sum_{i<j} n_{ij} \log (p_i + p_j) - \sum_i a_i \log p_i$. Then, if λ is the likelihood ratio

$-2\ell n\,\lambda = 2\,(\sum_{i<j} n_{ij})\ell n2 - 2B_1 \ell n10$, is distributed in the limit as a χ^2 with (t-1) degrees of freedom when all n_{ij}'s become large. For testing appropriateness of the model we can use the statistic

$$-2\ell n\,\lambda = 2\ell n10\,(\sum_{i \neq j} a_{ij} \log a_{ij} - \Sigma n_{ij} \log n_{ij} + B_1) \ldots\ldots (5.4)$$

which is distributed in the limit as a chi-square with (t-1) (t-2)/2 degrees of freedom when all n_{ij}'s become large. (5.4) is derived under the assumption

$P(X_i > X_j) = \bar{\pi}_{ij}$

and the statistic tests the hypothesis

H_o: $\bar{\pi}_{ij} = \pi_i/(\pi_i+\pi_j)$ $(i \neq j, i,j = 1,2,\ldots,t)$

against the alternative

H_a: $\bar{\pi}_{ij} \neq \pi_i/(\pi_i+\pi_j)$

for some i and j. The precision of the estimated P_i's can be computed by the formulae:

$$\lambda^{\circ}_{ii} = (t^2/4) \sum_{j \neq i} n_{ij} \quad (i=1,2,\ldots\ldots t)$$
$$\lambda^{\circ}_{ij} = -(t^2/4)\, n_{ij} \quad (i \neq j;\ i,j = 1,2,\ldots\ldots t) \qquad \ldots\ldots\ldots\ldots(5.5)$$

The superscript here indictes the null case when $\pi_i = \frac{1}{t}$ $(i=1,2\ldots t)$. Then the asymptotic variance-covariance matrix of the treatment ratings is given by the inverse matrix $[\lambda^{\circ}_{ij}]^{-1} = [\sigma^{\circ}_{ij}]$ $(i,j = 1,2,\ldots\ldots,(t-1))\ldots(5.6)$.

Efficiency in such experiments can be studied with the help of pairwise precision which gives the sample size N_{ij} which would be required in a two sample test in order to obtain $p_{ij} = \text{est}\,\pi_{ij} = p_i/(p_i+p_j)$ with the same precision as results from a paired comparison analysis. For the pair (i,j) it is given by

$$N_{ij} = 4/t^2(\sigma_{ii}{}^{\circ} + \sigma_{jj}{}^{\circ} - 2\sigma_{ij}{}^{\circ}) \;\ldots\ldots\; (5.7)$$

Step 3 - Compute the first estimates of p_i using (5.3). Substitute the estimates on the R.H.S. of (5.2) to get the subsequent estimates. The iteration is continued till two successive estimates coincide. To test whether the p_i's differ among themselves significantly, compute

$$-2\ell n\,\lambda = 2(\sum_{i<j} n_{ij})\,\ell n_2 - 2B_1 \ell n 10$$

where

$$B_1 = \sum_{i<j} n_{ij} \log\,(p_i+p_j) - \Sigma a_i \log p_i.$$

Test for significance of this using χ^2 table. For testing the appropriateness of the model compute the statistic (5.4) and test using chi-square tables. Compute the asymptotic variance-covariance matrix of p_i's using (5.5) and (5.6) and pairwise precisions by (5.7). The computations in the numerical example were shown in Step 1. The main statistical interpretation is that the food mixes differ significantly and that the model is appropriate. Interpretation in the subject matter area is that T2 is the more preferred food mix followed by T1, although they do not differ significantly. Comparison among the other foox mixes can be made in a similar manner. Pairwise precision for the pair (1,2) is 156 thereby meaning that in this experiment with 85 repetitions for this pair, the precision achieved is as good as in a single pair experiment with 156 repetitions. Thus paired comparison experiments improve the precision of treatment comparisons due to hidden replications. These results can be given to the food industry for practice.

In the next steps practicals can be given to the trainees and the record of class work kept. The trainees can do an experiment in sensory evaluation, analyse the data and write a report. This makes the training very comprehensive.

SOME REFLECTIONS ON TEACHING NON-STATISTICIANS

In this section we will refer to some dubious points. Should construction of designs be included in the curricula for scientists? The answer to this is possibly 'no'. For teaching collection and compilation of agricultural statistics are the teaching methods discussed in this article necessary? This also has a negative answer. This is not so for teaching research techniques in agricultural sciences and also for teaching statistics minor in universities. Does the training of non-statisticians to some extent bridge the gap between academic statisticians and practising statisticians? The answer to this is possibly 'yes' because some of the practising statisticians are non-statisticians. Should all the non-statisticians trained be assessed? In my opinion all need not be. There is also the fundamental question whether statistics should be taught to non-statisticians at all. The answer is that they should not be taught statistics to the detriment of statisticians.

CONCLUSIONS

Thus the method of teaching statistics to non-statisticians must be application-oriented. They should be trained as to how to present interpretative results. Examples must be given to demonstrate what can be inferred and what cannot be inferred from real data. The teaching method can be the same in each specialised area except in areas as collection of agricultural statistics where the contents are purely descriptive. A course in computer programming is a must for most types of non-statisticians. Field work involving independent designing of data collection and analysis is also a must for most non-statisticians.

For assessing agricultural scientists we have been using the credit-course system of America. The trainees are given quizzes, mid-terms and a final examination according to this system. Maximum weight is given for the final examination in the four point grading system adopted here. Further, there are comprehensive examinations and vivas in the relevant subjects as well as for the thesis work of non-statisticians. Similar methods are adopted for grading non-statisticians in Agricultural Universities and Indian Institutes of Technology. For Junior, Senior and in-service training courses of the IASRI, we adopt the external examination system of the Indian Universities including a "viva voce" examination. Students in these courses are awarded classes. For traineees of All India Services, the courses given are seminar type with neither assessment nor grading. For students joining on short term projects, the project reports are assessed. Over the years we have developed some specialised curricula for the non-statisticians. Teaching materials for such trainees have also been developed. The results of the consolidated work have been published as 'Statistical Methods for Agricultural Workers' and

'Handbooks of Practicals' apart from a number of applicational papers which appear from time to time in the Journal of the Indian Society of Agricultural Statistics.

References

[1] Cochran, W.G. (1953) Sampling Techniques. John Wiley & Sons Inc.

[2] Kempthorne, O. (1980) The Teaching of Statistics. The American Statistician. 34, No.1, pp.17-21.

[3] Mosteller, F. (1980) Classroom and Platform Performance. The American Statistician. 34, No.1, pp.11-16.

[4] Noether, G.E. (1980) The Role of Non-parametrics in Introductory Statistics Courses. The American Statistician, 34, No.1, pp.22-23.

[5] Sadasivan, G. (1966) Statistics and Agricultural Education. Journal of University Education, Delhi University, 5, No.1. pp.19-24.

[6] Tukey, J.W. (1980) We need both expository and confirmatory. The American Statistician, 34, No.1, pp.23-25.

[7] Davidson, R.R. and Beaver, R.J. (1977) On extending the Bradley-Terry model to incorporate within-pair order effects. Biometrics, 33, pp.693-702.

[8] Dykstra, O. Jr. (1960) Rank analysis of incomplete block designs; a method of paired comparisons employing unequal repetitions. Biometrics, 16, pp.176-188.

UTILITY THROUGH COLLABORATION: A PROBLEM ORIENTED APPROACH

D Taylor and G E Skipworth, Teesside Polytechnic

The teaching of statistics to non-statisticians is a well established practice. Indeed such is the success with which the argument for the relevancy of statistics has been put that the practice has been extended to include undergraduate and postgraduate courses not just for engineers, nor only geographers and sociologists, but even such inveterately literary characters as historians! While the latter might be considered an extreme case, there is a common problem to be found in all courses of this nature. Put simply, it consists of convincing students of the relevance and utility of statistics to their particular discipline which in turn requires the conversion of plausible theory into credible practice. The theoretical justification receives considerable attention, especially from the plethora of validating bodies to be found in certain sectors of higher education, but too little time has been given to the more mundane, but no less important, task of actually teaching a course in a manner that successfully converts theory into practice.

The most common practice found in the teaching of such courses is that of the 'invited alien'. Given the absence of staff with the necessary statistical knowledge in the host department (a situation which may be changing slowly), it is not surprising to find that the 'visiting statistician' is called in to do the job. All too often, however, he is left to work in isolation. Lacking the detailed knowledge of the subject that is the prime interest of the students, he becomes an outsider, a once (or twice) weekly performer with a display of statistical pyrotechnics that, depending upon the nature of the audience, arouses either admiration for their elegant irrelevance or deep scepticism for their inhuman manipulation of real facts! This difficulty is compounded by the lack of time commonly devoted to such courses. Despite the importance attached to statistics courses by host departments, there are always more important indigenous courses that crowd the timetable and restrict the visiting statistician to little more than a token appearance. In these circumstances it is not surprising to find that the course becomes a whistle stop tour of some of the better known statistical techniques. If it's week three it must be correlation! Faced with these practical difficulties, noble aspirations quickly turn to unsatisfactory practice. Unsatisfactory for both staff and student as a series of techniques, thinly disguised with examples from the host department's discipline but with limited relevance to the day-to-day work of the students, are presented in a manner that has more in common with forced feeding than genuine scholarship.

Although a general dissatisfaction with the practice of servicing work was an important element in the formulation of the particular course that is the main subject of this paper, there were also more specific considerations that were to the forefront in the construction and evaluation of the course. Indeed, we would argue that the emphasis must be on specific operations rather than on general crusades. It was clear from preliminary soundings that, while the staff mounting the M.A. in Local History were convinced of the usefulness of statistics to their area of study and hence of the value of an introductory course for the students, many prospective students would have no knowledge beyond basic arithmetical skills and, moreover, would exhibit apprehension if not antipathy to any course that involved statistics. In these circumstances it was felt that an introductory statistics course could only succeed if it were clearly demonstrated to the students that it was an integral part of the whole M.A. programme and relevant to work that was taking place elsewhere in the programme. To this end we sought to show that statistical skills were both a necessary and natural part of the historian's 'tool-bag' and that these skills could be applied to problems introduced elsewhere in the students' studies. Utility through collaboration became the motto of the course, which was not only conceived but also put into practice as a joint venture. Statistician and historian would be seen working together and in a way that was mutually beneficial.

It is now appropriate to consider in more detail the Introduction to Statistics course in the M.A. in Local History offered at Teesside Polytechnic. The first point to stress is that each session lasts approximately three hours. Although this is in part the outcome of the fact that the M.A. is a part-time evening only course, a deliberate decision was taken to utilise the maximum time period so that a self-contained and integrated topic could be presented in any one session. The second point to emphasise is that, with the exception of the introductory meeting, each session is devoted to a specific topic and, moreover, a topic that relates directly to other courses in the M.A. programme. Thus, for example, one session is given over to a consideration of the use of measures of association to throw light on the determinants of nineteenth century voting behaviour, a problem that is of major importance on another, essentially historical, course that is devoted to a consideration of regional patterns of political action.

However, to illustrate the approach used on the course we will consider another example in greater detail. The question of changing living standards is one that has generated much argument among historians. One part of this debate has centred on the measurement of real wages as a measure of material well-being. The technique most commonly, and all too often uncritically, used is that of index numbers. This, then, provides the general problem area for a session. The session is divided into three roughly equal parts, but throughout every session both members of staff - statistician and historian are present and will contribute, though the balance will vary from one sub-period to another. The opening sub-period is led by the historian who firstly outlines the general problem, the range of opinions on the subject and, given the nature of the M.A., the relevance of the local study to the wider issue.

As part of this preview the case for using statistical techniques is also presented. In this instance particular emphasis is placed on the need to generalise, the level of generalisation appropriate to the question under review and, particularly, the limitations of the concept of real wages in itself and in relation to the wider issue of living standards. His second, and equally important, task is to provide basic data and a critical commentary thereon. This last point is doubly important. From the historian's point of view the critical scrutiny of material is a central part of his practice. In this case particular attention is paid to the difficulties of assembling data relating to working class expenditure and consumption patterns that underlay any attempt to construct a cost of living index. Budgetary information for the late nineteenth century is limited. It is not possible to give details of year to year changes, nor is there specifically local material to work on. With regard to local prices the position is brighter. The local press as well as the minutes of the local Board of Guardians provide valuable material but even here there is the problem of using adjusted institutional prices to be recognised. The need to recognise the imperfections of the data, however, is no less relevant to the statistician. Although applied statisticians, rightly, emphasise the practical nature of their work, there is a tendency in teaching the subject to have perfect examples. Understandable though this is, there are certain dangers inherent in such an approach - most notably oversimplification. However, in this course all concerned are face-to-face with imperfect data from the outset. Moreover, the question is constantly posed: is the data of sufficient substance and quality to justify further analysis? Clearly, with the raising of such a question the expertise of the statistician is brought in to the discussion.

The second part of the session, predicated on the assumption that the reply to the above question will be, however heavily qualified, in the affirmative, is led by the statistician. In this he sets out the basic principles underlying, in this case, Index Numbers. This is done via a combination of commonsense and basic arithmetic rather than detailed mathematical proofs. Further he makes clear any requirements of the data demanded by a particular technique (something, clearly, that will vary from technique to technique). Finally, he draws attention to the limitations of the technique itself. All three aspects deserve further comment. First, the presentation of the technique is seen to be a middle way between one in which mathematical proof is seen to be a prerequisite, inappropriate given the background of the students on this particular course, and one in which students are given a 'recipe book' of formulae, pre-digested rules to be applied without thinking, inappropriate to a course at graduate or undergraduate level. Second, the emphasis on the data requirements from the statistician's viewpoint is felt to be a most necessary part given the propensity of many non-statisticians, and often those most enthusiastic about quantification, to apply statistical techniques in any circumstances. Finally, the consideration of the imperfections of statistical techniques is seen as a necessary antidote to the common belief among non-statisticians that statistics offer certitude in both method and result. As a result of this sceptical approach the earlier question asked of the data is supplemented by another: does the technique provide a satisfactory approach to the problem in hand? This brings the historian back into the discussion and underlines the interplay between the two disciplines.

The third and final sub-period involves a consideration and interpretation of the results obtained from the application of the technique considered in part two to the data introduced in part one. The students are not expected to work out the results for themselves but are presented with a combination of handouts and overhead projections. In the evaluation both statistician and historian contribute to an assessment of both the statistical and the historical significance of the results. It is also at this point that the students can make (and indeed are expected to make) a contribution to the general discussion. It will be readily appreciated that this part of the session, and the related follow-up essay, is both the culmination of what has gone before and the justification for the whole exercise.

Having outlined the approach adopted in the course it is necessary to see to what extent the theory has stood up in practice. So far there have been two intakes of students which provide the empirical basis of our evaluation. It was clear from the outset that the major problem would be one of weighting both within topics in any given year and between groups from year to year. The first intake, a dozen in size, was relatively homogenous in terms of ability, sympathetic to the aims of the course, and well motivated. One or two, significantly non-historians by first-degree, had an elementary knowledge of basic statistics but the majority could be considered, and certainly considered themselves to be effectively innumerate, at the outset at least! The first running of the course revealed two weaknesses. First, the balance between the three sub-periods was difficult to maintain. In particular the third sub-period tended to be squeezed in terms of time and this was further compounded by the fact that students were responding immediately and in some cases to historical problems that they were meeting for the first time. To some extent this was masked by the enthusiasm of the students; in another direction, it was also offset by the follow-up essay work, though it should be made clear that the essay work was based on only one of the topics, chosen by the student from the range offered on the course. The second weakness, and considered to be only minor in the light of the first running of the course, was the lack of any opportunity for the students to 'dirty their hands' albeit on a small scale. Interestingly, this shortcoming was regretted by the students as well as the staff. In the light of this assessment it was decided for the second cycle to retain the same basic format with two modifications: (i) the deletion of one topic and (ii) the introduction of substantial handouts relating to sub-period two. In these ways it was hoped to allow more time for discussion and thereby meeting our original aspirations.

The second intake, twenty in number, exhibited different characteristics. The range of ability in both general and particular terms was much greater than with the first intake. Equally significant, while there was little difference in terms of overall motivation, there was an element not so much of scepticism but of cynicism about the objectives of the course. While this was overcome for the most part, there remained the two weaknesses identified in the first cycle but intensified by the first characteristic of this group. Superficially, the problem of internal balance was solved in that more time was made available for discussion but a more serious problem emerged, namely bringing those at the lower end of the ability range up to the necessary level of understanding to participate fully and satisfactorily in discussion, and indeed to produce the required standard of written work. This led us to re-consider what had been seen as the secondary weakness from the first cycle, that is the lack of any opportunity for students to 'dirty their hands'. It now seems clear that such an opportunity would be beneficial not only to strong students but also to those weaker students who had struggled to grasp some of the basic concepts introduced in the course. Thus, the introduction of partly worked examples, offering the students the chance of applying by imitation the techniques introduced in the session, should give a firmer grasp of the subject to all students, but particularly to the less able and less sure. Finally, the introduction of a 'dirty hands' element will necessitate a carry-over from one week to the next. While this obviously destroys the unity of each session it is felt that this will be more than compensated for by the firmer grasp of the subject matter and by the greater opportunity for reflection on the results.

The emphasis in the preceding paragraphs has been on the shortcomings of the course in its two presentations. However, it would be wrong to give the impression that the approach has been unsuccessful in overall terms. Given the backgrounds of the students coming onto the course and the limited time at our disposal the results, as reflected in essay marks and examination performances, have been encouraging. Students with little or no formal knowledge of statistics at the start of the course have finished with not just a grasp of basic techniques but also an appreciation of how those techniques can be applied in a proper manner which will further their work as practising historians. Even in other, less successful cases it is by no means clear that the shortcomings are confined to the statistics course. There are, however, other less formal but no less important measures of success. First, student response - almost unanimously in the case of the first intake, and a clear majority in the second - was positive and encouraging, with many wishing to take the subject further. Second, several students, in the course of their M.A. dissertations, utilised both the specific knowledge gained from the course and the general expertise of the staff. Finally, and no whit less important, the staff response has been favourable! Both authors have found the course not simply novel (though in their direct experience it can certainly claim novelty) but both stimulating and satisfying. This is not to say that there have not been problems and periods of heart-searching but it is a measure of the overall success of the course that these very shortcomings have acted as a stimulus to further refinements and improvements to what is seen as a basically sound way of teaching statistics to non-statisticians. No longer need the statistician be an 'invited alien': genuine collaboration can bring mutual benefits to both staff and students and to the disciplines involved in such an integrated offering.

AN OPEN UNIVERSITY SERVICE COURSE

John Bibby, The Open University

INTRODUCTION

This paper describes an Open University course entitled:

MDST 242: Statistics in Society

This course will be available to Open University students for up to 10 years starting in February 1983. However, in addition to being anecdotal and specific, the paper also raises general issues and questions concerning the teaching of statistical methods to nonspecialists.

DEFINING THE CURRICULUM

In his introductory speech to this conference, Vic Barnett has already raised the question

WHY TEACH STATISTICS?

He has, indeed, also provided tentative answers. However, a full answer will depend upon our attitude to a further crucial question, namely:

WHAT IS STATISTICS?

This can be answered in several different ways. The conference has discussed various competing definitions of statistics, and not all teachers can afford the luxury of questioning the curriculum - this is often defined by our examining body, head of department, departmental tradition, or other institutional constraints. In such situations one may easily unquestioningly accept or rationalise the status quo, and avoid returning to first principles.

However, many teachers do have considerable latitude in determining their syllabus. Some few are prepared to stretch institutional constraints to their limit. This is certainly the case at The Open University - possibly excessively so. We have the power (and problems) associated with rethinking the curriculum from first principles. This power must of course be exercised responsibly, although it is not always easy to know what such responsibility implies, when traditions and technology in our subject are changing as rapidly as they are today.

CONSTRAINTS

Nevertheless, way back in 1979 when planning for the course MDST 242 started in earnest, we did face certain constraints. Paramount among these were the following:

(a) We were concerned with an individualised learning system - students would study alone, at homes scattered throughout the United Kingdom, with minimal face-to-face tuition. The teaching material had to respond to this.

(b) We were aiming at a course which was non-mathematical. "Statistical Literacy", "Education for Everyman" and "Citizen-Arming" are watchwords which describe the approach we were aiming at. Moreover, we wanted to attract students from all six Open University Faculties. Therefore our applications could not be too discipline-specific.

(c) Our resources were finite. That is not to say they were inconsiderable - we could count on 15-20 person-years of academic person-power, and also other Open University resources such as text, audio-tape, television programmes and computing. (However, we decided not to include any computing, for various reasons including obsolescence-phobia).

(d) The most important constraint however was student time. Our course should take around 160 hours of study time in total. This may be compared with a lecture course of 40-50 hours (assuming 2-3 hours of back-up work per hour of lecturing).

(e) Moreover, we were conscious (and still are) that this could be an extremely influential venture (for better or worse), for statistics teaching across far greater horizons than just the Open University. We have 100 or so tutors scattered across the country, and many of these also teach at other colleges and universities. Other institutions, we hope, will consider using our work. Also, in the ten years of presentation (1983-1992) our course would reach up to 20 thousand students - a figure which is several orders of magnitude bigger than many other statistical ventures. Thus we sought a reasonable compromise between the requirements of tradition and those of good statistical practice.

FIGURE 1: Tensions to be resolved

(1)	Statistics as an "exact science"	↔	Statistics as a "social product"
(2)	a few "case studies"	↔	many "examples"
(3)	descriptive/ exploratory statistics (EDA)	↔	inference

TENSIONS

In the course of preparing our material, many decisions have had to be made. However, three key areas may be identified as encapsulating dilemmas which many teachers must face. These are summarised as the "tensions" shown in Figure 1.

(a) First is the tension between statistics as an *"exact science"* (objective, rigorous, culture-free, technique-oriented) and statistics as a *"social product"* (produced as the outcome of human responses to a wide variety of conflict-laden situations).

In resolving this tension we resolved that we *did* wish to teach technique (this was not to be simply a "statistical appreciation" course), but we also wished to sensitise our students to the technical and non-technical limitations of technique.

(b) The second tension was between presenting a few "case studies" on the one hand or many different examples on the other hand. The first course of action would allow us to go into material in some depth. However, it might unwittingly provide hostages to fortune, in view of the wide variety of subject-backgrounds from which our students would be drawn. Also, the second option could be too disconnected.

Therefore we resolved to focus on a few "topic-areas", from each of which a variety of applications would be selected. The topic-areas should involve students as *citizens* (producers, consumers, parents, etc.) rather than as adherents of a particular discipline. Thus we focussed on the following four main topic-areas:

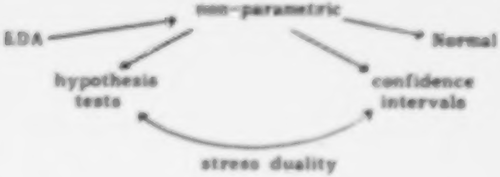
Economics
Education
Health and Medicine
Energy

Moreover we sought - somewhat optimistically as it turned out - to teach only techniques which "arose naturally" out of these topic-areas.

(c) The third tension referred to in Figure 1 relates to the balance between exploratory and data-reductive techniques on the one hand (à la Tukey/Ehrenberg) and classical inference on the other hand. This tension is certainly extremely topical now. However, we believe that EDA is no transient bandwagon. It reflects a long-standing unease concerning the role of probability-based methods. The use of these methods often carries a certain lack of conviction. Moreover, the philosophy and mathematics on which they are based can be obscure and are often positively intimidating.

However, they survive _faute de mieux_, and clearly are relevant in certain well-defined situations such as sampling and randomised experiments.

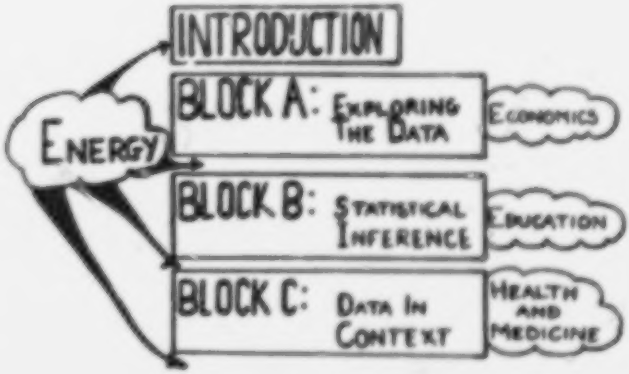
FIGURE 2: Our approach linking EDA with conventional inference



EDA is more generally applicable, but is less precise in the answers it gives. Thus, we sought a balance between exploration and inference. The balance we found is summarised in Figure 2: the course starts with EDA, and then approaches inferential statistics via a limited range of non-parametric methods.

We start with the sign test for the median, and its associated confidence interval. This has the advantage of mathematical simplicity - the pattern of Pascal's triangle is crucial here - which allows us to concentrate on teaching statistical points such as the duality between hypothesis testing and confidence intervals. The Normal distribution arises via the large-sample properties of primarily non-parametric methods - we are currently planning a television programme on this.

FIGURE 3: Structure of MDST 242



OUTLINE OF COURSE

Figure 3 outlines the course structure. Further details of structure and content appear in a Course Description available from the author. However, the main features are the following:

The course has 3 main Blocks, each consisting of 4 units. These may be summarised as follows:

Block A covers EDA, focussing on topics from economics (prices, wages etc.). This block also covers line-fitting using Tukey's resistant method, price indices and survey methods.

Block B introduces inference, with educational examples providing the vehicle. We focus on the limited menu of techniques shown in Figure 4.

Block C consolidates the techniques of exploration and inference. It also discusses notions of experimentation (drug-testing) and public policy-making (smoking).

FIGURE 4: "Menu" showing main inferential techniques

MDST 242

MENU

Hors d'oeuvre

Triangle de Pascal

Starter

Sign test/confidence interval for median

Main Course

Morceau de Mann-Whitney
French fried Wilcoxons
Crême a la Chisquared

Dessert

z (No t)

Notes:

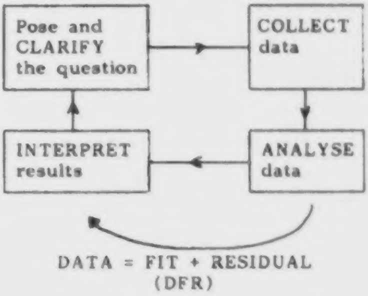
(1) Each confidence interval is served beside along and between the associated hypothesis toast, together with a strong gravy of duality.

(2) No one-tailed beasties.

(3) 5% significance only

Interleaving the 3 main blocks are 4 further units on energy statistics, using especially some of the work of the Energy Research Group at the Open University. These units are primarily "stock-taking" units : the student watches 2 television programmes, submits assignments, draws breath, and proceeds to the next block.

FIGURE 5: The model-building loop



STATISTICS AS MODEL-BUILDING

The "model-building loop" shown in Figure 5 is fundamental to the portrait of statistics which this course will paint. The loop has many directions of relevance:

- it ties in with the "Data = Fit + Residual" (DFR) equation of exploratory data analysis
- it illuminates the non-statistical problems associated with hypothesis-formulation
- it emphasises the iterative genesis and tentative status of statistical knowledge, thus supporting the notion of "statistics as a social product". (The debate in the sociology of science between Popper and Kuhn has many interesting parallels here).

The model-building aspect of statistics has scarcely been emphasised at this conference. This is not surprising, because the model-building viewpoint can query the very rationale of the conference's existence. It raises the question of whether statistics should be separated from mathematics at all. Applied mathematics properly defined is little more than deterministic model-building. By contrast, statistics is model-building done properly i.e. stochastically. This is a viewpoint which MDST 242 aims to emphasise.

However, while many courses concentrate on the COLLECT and ANALYSE boxes of the model-building diagram, our course (and in particular the television programmes) go to great lengths to illustrate the INTERPRET and CLARIFY boxes. Thus in TV3 (which was shown at the conference), fully half the programme is concerned with setting up the rationale for a line-fitting strategy in one particular context (the installation of cavity wall insulation). By contrast, the technique of line-fitting is correspondingly underplayed. (We do, however, find audio tape a very suitable medium for the teaching of techniques: it allows the student to stop and replay, self-assessment exercises can be interspersed; also technical details can be given a suitably low profile).

THINGS WE DID NOT INCLUDE (BUT MIGHT HAVE DONE)

All honest teachers learn from experience. Speaking personally, the experience of preparing MDST 242 has taught me many things. Predominant among these has been the joy and excitement of working with a group of lively-minded colleagues who are primarily concerned with applications, rather than theory. Despite occasional frustrations, preparing this course has been fun : and fun is a factor which should not be underplayed in any statistics course. (I only hope that the students enjoy the course as much as I have done; and also that they learn something from it!).

However, if given the opportunity to prepare the course again, there are several alternative possibilities which I would want, at the very least, to consider seriously.

(a) Statistics as a Social Product

This concept was referred to briefly in the section "Tensions". However, there remains an insidious force within our subject which pushes relentlessly towards technique. The tendency is enhanced by the fact that "statistics as a social product" remains an amorphous and ill-formulated concept : this is seen as weakness in a world where precision is a sign of strength. Nevertheless, I would wish that we had stood firmer against the tide of technique; although I am unsure whether a firmer resolution would actually increase our chance of success.

(b) Inequality as a Topic

Our choice of "Education" as a topic-area may have been misplaced, although undoubtedly it will prove to be popular with Open University students, many of whom are teachers. "Inequality" may have been a better choice - indeed inequality arises naturally in economics, in education, and also in health. However, a more explicit recognition of inequality as a theme may have been preferable. (Indeed, inequality and poverty are world-pervasive issues - a point which has not been lost on us concerning the world-wide marketability of our course, especially in developing countries).

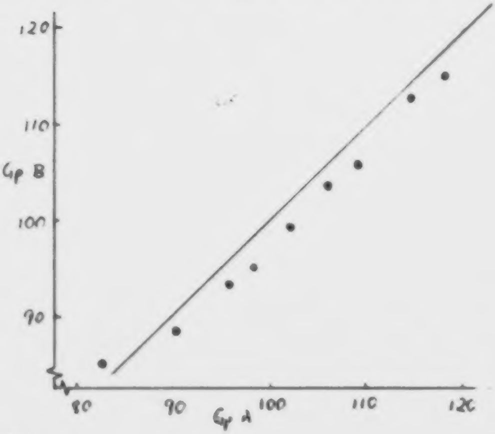
(c) Median Polishing

This technique, introduced by Tukey, is a non-parametric counterpart of analysis of variance. However, it is much, much simpler. The analysis of a 5x3 table can even be used to introduce the notion of median. Our early drafts proposed this, but unfortunately median polishing was subsequently winkled out by pressures of time and space, as well as by the feeling that the technique is as yet insufficiently used and understood.

FIGURE 6: Back-to-back stemplot of test scores achieved by members of two groups

GROUP A		GROUP B
	7	5
32	8	58
950	9	358
9621	10	34
974	11	35
	12	6

FIGURE 7: Quantile plot based on Figure 6



(d) Quantile plots

This method of comparing two batches is informal, informative, and relatively easy to use. As an example, the data on test scores presented by Professor Noether at the conference leads to the back-to-back stemplot shown in Figure 6, and to the quantile plot shown in Figure 7. If the two batches were identical, then each point would lie exactly on the 45° line. All other patterns are to be related to this bench mark or "FIT". Common patterns are the following:

- shift of 45° line : indicates difference in level
- tilt of 45° : indicates difference in scale
- "tailing off" : indicates something about the tails, outliers perhaps.

The plot shown in Figure 7 seems to indicate a difference in level between the two batches (Batch A greater than Batch B), with a slight reversal in the left-hand tail. However, there are problems associated with the definition of quantiles in small discrete batches, and further research may be necessary on the real-life properties of quantile plots before the technique is suitable for introduction into elementary statistics courses.

(e) Use of stemplots for easing calculations

Apart from its use in calculating order statistics, the elementary grouping provided by stemplots also provides convenient "short cuts" for calculating several standard statistics, in particular the mean, standard deviation and Mann-Whitney U statistic. The short cuts are similar to those provided by the "use of working mean" methods and are particularly useful when calculators are not available.

FIGURE 8: Use of stemplot for calculating the mean

s_i		f_i	$s_i f_i$	t_i
7	5	1	7	5
8	5 8	2	16	13
9	3 5 8	3	27	16
10	3 4	2	20	7
11	3 5	2	22	8
12	6	1	12	6
			104	55

Note: f_i represents the frequency at each level of the stem, and t_i represents the total of the leafs. (For further details see text).

(i) Calculating the mean

The mean is easily calculated from the stemplot shown in Figure 8, without using a calculator. The stemplot decomposition of a value x_i is

$$x_i = 10s_i + \ell_i \tag{*}$$

where s_i is the "stem value" of x_i and ℓ_i is its "leaf value".

In Figure 8, for instance, the stem value of 75 is 7, and the leaf value is 5. The rest of the calculation is now evident. We have

$$\Sigma x_i = 10\ \Sigma s_i + \Sigma \ell_i .$$

Now if f_i represents the frequency at level i of the stem, and if t_i represents the total of the leaves at this level, then

$$\Sigma\ s_i = \Sigma\ s_i f_i$$

and

$$\Sigma\ \ell_i = \Sigma t_i.$$

(Some liberties have been taken with notion here, in the interests of clarity and simplicity).

Therefore

$$\Sigma\ x_i = 10\ \Sigma\ s_i f_i + \Sigma\ t_i.$$

The righthand side involves fewer and simpler additions than the lefthand side. The data in Figure 8 gives

$$\Sigma\ x_i = (10 \times 104) + 55 = 1095,$$

a value which can easily be checked using conventional methods. The mean is now easily calculated.

(ii) Calculating the variance

The stemplot decomposition described above leads to several alternative methods of calculating the variance. The most obvious of these is the following:

From (*) we deduce that

$$x_i^2 = 100 s_i^2 + 20 s_i \ell_i + \ell_i^2.$$

Therefore $\Sigma x_i^2 = 100\ \Sigma\ f_i s_i^2 + 20\ \Sigma s_i t_i + \Sigma \ell_i^2.$

FIGURE 9: Calculation of variance based on Figure 8

s_i	f_i	t_i	$f_i s_i^2$	$s_i t_i$	$\Sigma \ell_i^2$
7	1	5	49	35	25
8	2	13	128	104	89
9	3	16	243	144	98
10	2	7	200	70	25
11	2	8	242	88	34
12	1	6	144	72	36
			1006	513	307

A worked example is shown in Figure 9. The first three columns are taken from Figure 8. Using the above formula gives

$$\Sigma x_i^2 = (100 \times 1006) + (20 \times 513) + 307 = \underline{111,167}.$$

The final term, involving t^2, is relatively unimportant in this calculation, and may even be omitted without affecting the result unduly.

(iii) Calculating the Mann-Whitney statistic

The Mann-Whitney statistic U is used for comparing two groups. It calculates for each member of one group the number of members of the other group which are less than it; U is then the sum of these numbers.

The stemplot provides an easy way of calculating U, and even easier ways of calculating upper and lower bounds. The method is illustrated in Figure 10.

FIGURE 10: Calculation of U based on Figure 6

Stem value	Group A cum.freq. (d)	Group B freq. (f)	fd	fd^{-1}	Corrections
7	-	1	-	-	-
8	2	2	4	-	4
9	5	3	15	6	4h
10	9	2	18	10	4
11	12	2	24	18	1
12	12	1	12	12	-
			73	46	13h

Note: "h" means "half"

The approximation with the clearest rationale is shown in the column headed fd^{-1}. Here each frequency from Group B is multiplied by the number of observations in Group A which are at a lower value on the stem (i.e. higher up the stemplot, because low values are at the top of the stemplot). The total of this column represents a lower bound on (i.e. is less than) the required value U.

An upper bound can be calculated as the total of the column headed fd. Alternatively the values fd^{-1} can be "corrected" by looking at the individual leaves at each level on the stem, one at a time. By referring back to Figure 6 the corrections shown in Figure 10 may be derived. Adding in these corrections we find the value of U is 59h or 59½. ("h" is a convenient shorthand for "half". It is simpler, quicker to write, and less dominating than "½").

A very close approximation to the value of U can be obtained by averaging the totals obtained in the columns headed fd and fd^{-1} (7 and 46 in this case).

CONCLUSION

Any course which is innovative is bound to have weaknesses as well as strengths. We hope to monitor feedback from students and tutors on our course, in order to improve its presentation in subsequent years.

However, it is also to be hoped that some ideas may be found here which merit application by other teachers from outside The Open University. The author would be very interested to hear of such applications, and of the results.

THE TRAINING OF SERVICE TEACHERS

R W Madsen, University of Missouri

Those of us who are attending this International Conference on Teaching Statistics no doubt have as great or greater interest in teaching than do many of our colleagues. Certainly as a group we have many years of teaching experience among us and many of us, over the years, have developed good teaching skills and techniques. I believe that most of us realise that we can still improve our teaching and that is one reason that we have come to this conference: so that we can learn some teaching ideas that are new (at least to us).

How have we learned our own teaching skills? Most of us have learned them from the "best teacher". At least we can say that we have if the cliche that "experience is the best teacher" is true! It is a well-known fact that, at least in American colleges and universities, most professors are highly trained in the subject matter of their chosen discipline but receive little or no formal training in how to teach that subject matter. These professors (including most of us) have learned teaching by experience and from informal observation of our own professors when we were students. This is quite a contrast to the training received by elementary and secondary school teachers who are required to take 40 or more semester hours of course work in areas such as Oral Communication, Psychology of learning, Psychological Development of the Adolescent, Education of Exception Student, etc. Many of us have in fact learned over the years to become reasonably good teachers, but it has taken more time than it might have had we received some formal training.

If it is true that formal training in teaching would be beneficial for individuals who have completed Ph.D. programs and who are attempting to teach at a college level, then such must be equally true for individuals who are only beginning their graduate education. In many American universities introductory level courses in many disciplines are taught at least in part by graduate teaching assistants. It is the purpose of this workshop to discuss together how these teaching assistants might be trained to teach.

Most of the teaching done by graduate teaching assistants is in service courses. Bancroft (1972) described service courses as "courses which are provided by one or more departments for students majoring in a different department or departments". Much has been written on the mode of teaching service courses. Fiscal restrictions generally dictate one of two models: large lectures taught to 300-500+ students by an experienced instructor (Sherr, 1981) with perhaps recitation-discussion sections of 30-50 students conducted by teaching assistants or smaller lectures taught to 30-50 students by teaching assistants. (Most of us would prefer to have smaller classes to allow for student interaction and to have these classes taught by experienced faculty but that is, unfortunately, not a fiscally admissible option!). The teaching of service courses, whether by teaching assistants or regular faculty, is often more difficult than teaching courses to our own majors because many of the students in the class are "required" to take the course and are not taking it because they want to. This difficulty makes it even more important to have trained mature instructors. (See Federer, 1980).

Havign discussed the need for the training of teachers of service courses, especially when these teachers are graduate teaching assistants, we might address the question of who should be involved in this training. The following quotation may give some idea: "Can one blind man be guide to another? Will they not both fall into the ditch? A pupil is not superior to his teacher; but everyone, when his training is complete, will reach his teacher's level..." (Luke 6: 39-40 NEB). Using this as a guide, we should have our best and most interested teachers lead the training of new teachers.

We have discussed the "why" and "who" of the training of teachers and now turn our attention to the question of "when" and "how". To this end we will first discuss what has been done in the Statistics Department at the University of Missouri-Columbia. We will also freely admit that this program of training, which has evolved over a period of six years, has not nearly achieved perfection. Even if it were a perfect program for us, it would not be a perfect program for you. It is our desire to interact with each other so that we might each find ways of making our training programs more effective. What follows is an outline of topics we have discussed with new teaching assistants and abbreviated versions of handouts given to them. In many cases questions are raised on the handouts, but the answers are left for the teaching assistants to fill in based on discussions within the group. We should also add that we have a small department (about 10 full time faculty) so we have had as few as three and only as many as eight in a group at a time. The amount of time devoted to each session is 30-50 minutes.

Session I: Keys to Effective Teaching

This session (as well as Sessions II and III) are conducted prior to the first day of class. We want the teaching assistant to think about their new role as "teacher" rather than "student".

Handout I

1. Understand what it means to "teach".
    (a) To show how to do something.
    (b) To provide with knowledge and insight.
2. Know the material to be taught.
    (a) Know what will be taught in the entire course.
    (b) Know answers to problems assigned.
3. Know your students.
    (a) Know the majors so that you can use applications of interest.
    (b) Know as many of the students as possible by name.
4. You must teach so that the students can learn.
    (a) You must do your part and encourage the student to do his part.
    (b) Motivate the student by showing the need for and applicability of probability and statistics.
    (c) Students can be encouraged to learn, but cannot be forced to learn.
    (d) Try to view your teaching from your student's point of view.
5. Involve the students in the learning process.
    (a) Ask them questions.
    (b) Allow them to ask questions.
    (c) Have them participate in generation of data and examples.

Session II: First Day Tips

After more than 14 years of teaching I still get nervous on the first day of class. In addition to new faces, the first day activities are quite different than on other days. Students may be in the wrong class, have the wrong text, etc. It is helpful for the new teachers to know explicitly things that they should or shouldn't do.

Handout II

1. Expect to be somewhat nervous, especially if this is your first day of teaching.
2. Locate the room in which you will teach in advance.
3. Begin and end this and all classes on time (ideally within 1 minute of the scheduled time.
4. Give information to the students.
    (a) Class name, number, and section.
    (b) Prerequisites.
    (c) Your name, office location, phone.

5. Get information from the students.
   - (a) Name (to check against class roll).
   - (b) When prerequisite was taken.
6. Give information about course mechanics.
   - (a) Homework policy.
   - (b) Exam schedule.
   - (c) Grading
7. Refer any out-of-the-ordinary question to a faculty advisor.
8. Be prepared to give a regular lecture on the first day.

Session III: Lecture Preparation

One of the best articles I have seen written on the teaching of statistics is by Mosteller (1980). He writes "... I plan to bring out a few ideas that I have found useful in teaching. Although no two teachers are alike, I have often found devices used by others adaptable to my own style... I shall speak as if to new teachers eager for my advice". It is in this spirit that I present my ideas on lecture preparation to new teaching assistants.

Handout III

1. Prepare in advance.
   - (a) Do not wait until the hour before class.
   - (b) Prepare about 1/3 to 2/3's of the next lecture in addition to preparation for the present one.
   - (c) News articles, examples from the literature, etc. should be filed when found for future use.
2. Follow a schedule/outline.
   - (a) Use departmental guidelines.
   - (b) Construct your own.
3. For each main concept use a three step method.
   - (a) Motivate
     - (i) Describe a realistic problem which needs to be solved. Problem used should be of interest to your students.
     - (ii) Clearly state basic assumptions including simplifying assumptions.
     - (iii) Leave the problem unsolved at this point.
   - (b) Present Theory
     - (i) Make transition from the particular motivating problem to a more general setting.
     - (ii) Use notation consistent with the text. If the notation is at variance with standard notation, make note of the fact.
     - (iii) Use appropriate mathematical level. You should neither oversimplify nor be overly rigorous.

(c) Illustrate the Theory

  (i) Solve the problem raised in part 3(a). Identify parameters. Explain where and why simplification was necessary.

  (ii) Illustrate mechanics. Work out mechanics in some cases and/or give students an opportunity to try on their own. (Be careful to make good use of class time on this. Calculations should take only 1 to 2 minutes or less).

  (iii) Try to provide insight by using a word problem. Explain how you check conditions to know that the theory applies.

4. Prepare notes of adequate detail.

  (a) Write enough so that you can follow your own work.

  (b) Work out numerical answers in advance.

5. BE ORIGINAL.

  (a) Do not use examples directly from the text (except in special circumstances

  (b) Modify your own notes from semester to semester.

Session IV: Lecture Delivery

On the assumption that it is not possible to assimilate a great deal of information in a short period of time, after the first three sessions we meet once a week for a teaching seminar. One of the important topics to cover early is that of lecture delivery. Mosteller (1980) also makes some pertinent comments on delivery.

Handout IV

1. Spoken Communication

  (a) Speak to the class. Make eye contact with each person.

  (b) Speak clearly. If you have an accent, be aware of it and make extra efforts to overcome it.

  (c) Speak loudly enough so that you can easily be heard in the far corner of the room.

  (d) Avoid a monotone voice.

2. Written Communication

  (a) Write legibly. Print if your writing tends to be sloppy. Also write large enough so that writing can be read from the back of the room.

  (b) Be careful about location. Do not write too low on the board or too far from the middle.

  (c) Write down all important information. Most students will write in their notes only that material which is written on the board.

  (d) Make use of graphs and diagrams. Use colored chalk as appropriate.

3. Nonverbal Communication

  (a) Your own enthusiasm about teaching and about the subject matter is important. It has a strong effect on student attitude toward the course.

(b) How you say things, e.g. your tone of voice and attitude, may be as important as what you say to the student.

Session V: Exam Construction and Grading

There are two schools of thought on the administration of examinations for multisection courses. One is to have a common departmental exam for all sections and the other is to have different exams made up and administered by the individual teachers. In our department we do some of both, so we want to give our new teachers some help in knowing how to construct exams. Since good multiple choice exams are harder to construct (although easier to grade) we have concetnrated on the written exam format.

Handout V

1. First steps
    - (a) Determine in advance the number of exams and the relative weights of the exams.
    - (b) Announce the date of the exams at least one week in advance.
    - (c) Tell the students what material they are responsible to know.
2. Characteristics of "good" and "bad" exams.
    - (a) As perceived by the students.
    - (b) As perceived by the instructor.
3. Exam construction
    - (a) Make a list of the various topics and concepts to be covered.
    - (b) Write the exam remembering the following:
        - (i) Give explicit instructions.
        - (ii) Cover as many topics as possible within time constraints.
        - (iii) Use some of the following types of questions.
            1. Definitions.
            2. Routine mechanical problems.
            3. Application or word problems.
        - (iv) Questions should have the property of
            1. Being legible.
            2. Having different parts of questions independent.
            3. Being unambiguous.
            4. Giving point values.
            5. Having appropriate mathematical level.
4. Proctoring of exams.
    - (a) Discourage cheating by setting up appropriate physical conditions.
    - (b) Discourage cheating by your physical presence.

5. Grading exams.
   - (a) Promptly.
   - (b) Give partial credit consistently.
   - (c) Show where errors are made.

6. Learning from the exams.
   - (a) You should make note of which questions are too hard or too easy.
   - (b) Go over the exam questions, especially the difficult ones, so that the students can learn from their mistakes.

Session VI: Involving Students in the Learning Process

One of the chief advantages of a small number of students in a class is that there is opportunity for teacher-student interaction. If we do not make use of these opportunities we are, I believe, doing the students a disservice. Mosteller (1980) states "Gathering relevant data in class and then analyzing them with the help of the class gets the students involved because we are analyzing their data and they can contribute to the analysis". Hansen (1980) makes a similar statement: "Students are more interested and attentive when they are involved in the class. In applied statistics class[es] which discuss data analysis, a physical demonstration is one good way in involving the students".

Handout VI

1. Allow and encourage students to ask questions:
   - (a) On homework or previous lectures.
   - (b) On the current lecture (at appropriate points in time).

2. Ask the students questions.
   - (a) Some questions may be rhetorical.
   - (b) For non-rhetorical questions allow the students time to think.
   - (c) Have students work out simple computational practice problems.

3. Have students participate in generation of data and conduct simple experiments.
   - (a) Let individual students perform simple experiments. (Coins, dice, cards, etc.).
   - (b) Let the entire class participate in some experiments and then summarize the results. This approach can readily be used in studying binomial, negative binomial, Poisson, or uniform distributions as well as for the Central Limit Theorem and random sampling.
   - (c) Use the class as a source of data.
     - (i) Can be useful for descriptive statistics, study of proportions (binomial), distribution of random variable, etc.
     - (ii) Has positive effect on class attitude in that people (frequently) like to talk about themselves and are curious about others in the class. (E.g. question about family size, state in which they were born, birthday problem, etc.).

(iii) Can be negative if you ask the wrong question. (Don't ask about low GPA's, divorced or alcoholic parents, or other potentially embarrassing questions).

Session VII. Use of Teaching Aids

The topic of teaching aids or teaching materials receives much attention. Several sessions at this conference have been devoted to the topic, the American Statistician has a section entitled "Accent on Teaching Materials", and the journal Teaching Statistics often contains articles on teaching aids. New teachers of statistics often need to be encouraged to use teaching aids and should be told how to incorporate them into a lecture.

Handout VII

1. **What are teaching aids?**
2. **What are some purposes of teaching aids?**
   - (a) Gain and hold attention.
   - (b) Add interest.
   - (c) Increase retention of knowledge through association.
   - (d) Show applicability to real world situations.
3. **What are sources of teaching aids?**
   - (a) Commercially prepared aids.
   - (b) Adult or children's games.
   - (c) Newspapers, magazines.
   - (d) Homemade materials.
4. Types of teaching aids.
   - (a) Visual aids
     - (i) Dice, cards, coins, marbles, etc.
     - (ii) Colored chalk.
     - (iii) Cartoons.
     - (iv) Packages of products, examples of defectives.
   - (b) Audio aids. Tape recordings of sports, speeches, news, weather, etc.
   - (c) Printed aids.
     - (i) Newspaper and magazine articles. (Separate handout on these).
     - (ii) Mail advertisements.
     - (iii) Books - especially from courses students will take (or have taken).
   - (d) Projected aids.
     - (i) Overhead projector with transparencies.
     - (ii) Slides, movies.
5. **When should teaching aids be used.**
   - (a) **Whenever appropriate, but most frequently as part of introductory motivating example.**

(b) As often as possible. Most people err in using too few rather than too many teaching aids.

Session VIII. Teaching Aids - Newspapers and Magazines

One of the most convincing ways of showing students the applicability of statistics to everyday life is to use common everyday sources. Among the most common everyday sources are newspapers and magazines. In view of the importance of these items as teaching aids, a separate session is devoted to them.

Handout VIII

1. How can suitable material be found?
    (a) Purposefully.
    (b) Serendipitously.
2. Where should materials be kept?
    (a) Specific files can be made.
    (b) Clippings can be placed in the text at the appropriate place.
    (c) They should not be left in a newspaper or magazine to be cut out "later".
3. Can material be "re-used"?
    (a) Not in the case of dated or easily replaced materials.
    (b) Possibly in the case of "classic" examples.
4. For what subject matter will these aids be most useful?
    (a) Probability: precipitation probabilities; odds for sports events:
    (b) Misuses, abuses of statistics: Proper use of mean vs median; incorrect or misleading labels on graphs (especially in advertisements)
    (c) Descriptive Statistics: graphs, tables, histograms, pie charts, etc. (especially abundant in news magazines).
    (d) Estimation (point estimtes, interval estimates, question of sample size): Use proportiosn given in Gallop Poll etc. (Not strictly random, but probably representative). Also point out difficulties with "man-in-the-street" polls, etc.
    (e) Hypothesis testing (especially about proportions): Same as (d).

Session IX. Student Evaluations

One of the most controversial topics among educators is that of the use and value of student evaluations. Rencher (1979) states that "Widespread disagreement among faculty members about the merits of student ratings would seem to point up the fact that the concept of teaching effectiveness is a complex phenomenon to evaluate". I believe that part of the reason for the controversy is the fact that administrators sometimes use these evaluations, to greater or lesser degrees, in evaluating teaching. Our purpose is not to discuss the appropriateness of such use. I do believe, however, that individuals who desire to improve their teaching can make use of student evaluations for self-improvement. I encourage our teaching assistants to view them primarily in this light.

Handout IX

1. Discussion of validity of student evaluations.
2. Discussion of factors which might affect responses.
    - (a) Time of class.
    - (b) Required or elective course.
    - (c) Student grades.
3. A look at the evaluation form.
4. Use of evaluations.
    - (a) By faculty advisor.
    - (b) By teacher.
        - (i) Trimmed mean approach.
        - (ii) Mechanics.
        - (iii) Attitudes.
        - (iv) Language or accent problems.

References

[1] Bancroft, T.A. (1972) On Teaching of Service Courses in Statistics, The American Statistician, 26, 14-16.

[2] Federer, Walter T. (1980) Pressures For and Against Effective Teaching, Proc.Sec. on Statistical Education, American Statistical Association, 28-30.

[3] Hansen, James L. (1980) Using Physical Demonstrations When Teaching Data Analysis, Proc.Sec. on Statistical Education, American Statistical Association, 61-62.

[4] Mosteller, Frederick (1980) Classroom and Platform Performance, The American Statistician, 34, 11-17.

[5] Rencher, Alvin C. (1979) An Analysis of Student Evaluation of Instruction and Relationships Among Salary, Chairman's Ratings, and Student Ratings of Faculty, Proc.Sec. on Statistical Education, American Statistical Association, 1-5.

[6] Sherr, Lawrence A. (1981) Teaching the Introductory Course to Large Classes, Proc.Sec. on Statistical Education, American Statistical Association, 8-12.

TOPIC 5: DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL

ORGANISER: MR P HOLMES

DEVELOPMENT OF TEACHING MATERIALS AT SCHOOL LEVEL

A R Hoffer, University of Oregon

It is clear to all of those who are attending this conference that statistics should be an important component in the education of all people and that statistics education should begin with young children and be carried on continuously throughout the years of formal education. However, a rather large percentage of teachers, especially at the school level, have not themselves experienced an adequate statistics education and hence lack knowledge about statistical topics per se as well as what types of statistical activities would be appropriate for their students and how to integrate such activities into their course of study. It is necessary to attack this cycle of ignorance from several directions: (1) the initial training of teachers should have an adequate statistical component; (2) there should be a variety of inservice programs for practising teachers at colleges, within school systems, and for teacher self-study; (3) curriculum developers and directors should more enthusiastically incorporate adequate statistical activities in their programs and courses of study; (4) there should be more active research into questions of statistical didactics; and (5) there should be developed more effective statistical lessons and learning sequences for students. Concerning the item of research most of the available studies deal with probability rather than other aspects of statistics.

There is a nice analogy between statistics learning and computer science learning. It is insufficient just to talk about all the potential uses of computers, and it is not enough to use only pre-prepared programs. Students need to write their own programs as well as to pose and solve their own problems with computers. Likewise one needs to "do" statistics. Carrying the analogy further, there are people, including middle school teachers, who are afraid of computers just as there are those who are afraid of statistics, in so far as they understand the topic. Hence to improve statistics education at the school level we must work to help teachers gain a positive attitude toward the subject as well as to increase their knowledge.

There are two more obstacles of which one must be aware as we develop statistics teaching materials. First, we must realize that statistics makes demands for sophisticated reasoning processes to which most people are unaccustomed. The usual activities at the school level deal with arithmetic, algebraic and geometric computations. Hence it is a fairly simple matter to instruct students to distinguish between different types of averages or compute means and so forth. It is more difficult to develop skills in collecting and organizing data, analyzing probabilities, testing hypotheses, and more generally thinking critically and being comfortable with uncertainty. A second obstacle deals with the heavy demands that are placed on elementary school teachers. They are usually responsible for covering many subjects besides mathematics and statistics. (As we note below this can be turned into an advantage). They must find time and energy to devote, for example, to the daily situations of students with varying abilities, to the demands of parents and school administrators, to the results of research, new learning theories and teaching strategies, and to the new teaching materials in the various disciplines. It is folly to expect that simply by developing more teaching materials at the school level will we be assured that more statistics learning will take place. The materials must be of the type that are readily acceptable by teachers. Indeed a proof of this situation may be evidenced by the availability over the past twenty years of numerous statistics teaching references with little or no change in school programs.

Statistics is a dynamic endeavor, forever changing. By their very nature statistical worksheets and textbooks with pre-prepared data are outmoded by the time they are printed. The real import of statistics learning rests in the relevance to learners as to how statistics is used to affect their lives and how they can use statistics to more fully understand phenomena. This points out a need then for at least two types of teaching materials: one type deals with the development of background knowledge and basic statistical skills while the second deals with engendering in the learner the abilities to recognize in ill defined situations how statistics can be used to obtain more information and to help in making decisions.

In the school year 1976-1977 we brought together at the University of Oregon a team of people to develop a resource on statistics for teachers of middle school children (ages 10-13 years). The project writing staff consisted of two middle school teachers, a high school teacher, a statistician, a mathematics educator and a mathematician. The resource, entitled Statistics and Information Organization [see Hoffer, 1977] contains background material for the teachers on statistical content and didactics along with an extensive collection of suggested classroom activities. Although the staff collaborated on all aspects of the writing, for specific sections each person had responsibilities which included an analysis of existing materials and publications, outlines, and writing of initial drafts. The resource consists of the following components:

Content for Teachers
Didactics
Teaching Emphases
Classroom Materials
Glossary
Annotated Bibliography
Selected Answers

The Content for Teachers section provides teachers with background information on statistics and probability. The subsections deal with gathering data, tables, graphs, scatter diagrams, averages, deviations, sampling, experimental probability, probability with models, counting techniques, probability with counting, and inferential statistics. There is also an appendix with 22 computer programs. There are numerous references here to other parts of the resource especially the classroom materials that serve as examples of the ideas as well as providing the teachers with connections to potential work for their students.

The Didactics section provides teachers with information on learning theories, teaching techniques, diagnosis and evaluations, goals and objectives. There are separate articles on components of instruction, classroom management, and statistics and probability learning. The latter includes a survey of the available research studies. Once again these articles contain several references to other components in the resource, especially the classroom materials.

The Teaching Emphases section contains a collection of processes, approaches, and aids that are emphasized throughout the resource. The subsections deal with critical thinking, decision making, problem solving, models and simulations, calculators and computers, and laboratory approaches. As the name implies these are aspects that should be emphasized as we teach statistics in addition to specific subject matter topics. There are cross references in the teaching emphases sections to the content, didactics and classroom materials sections of the resource.

The Classroom Materials section deals with concepts, skills and applications and takes up over half of the resource. The sixteen subsections here cover many of the usual topics including gathering data, scatter diagrams, sampling and monte carlo methods. The areas of application deal with business and commerce, environment, health and medicine, people and culture, and recreation. Before each subsection there are teacher commentaries which give a rationale for teaching the topic; suggest alternate ways to introduce or develop topics; suggest ways to involve students; highlight the classroom pages; and give more ideas that interrelate with the content, didactics and teaching emphases sections. The classroom materials are of two basic types: one intended for direct student use and the other as suggested activities directed by the teachers. In both cases there are short lessons as well as sequences of integrated lessons. Each subsection contains worksheets that teachers can duplicate for class use; activity cards that may be laminated; transparency masters; games for small group or class use; teacher directed activities; teacher ideas that can be used to develop other activities; and teacher pages that give teachers mathematical background for specific student pages.

To round out the resource there is a Glossary that gives informal definitions and an Annotated Bibliography that provides teachers with sources for additional ideas and references.

As noted above, the various components of the resource are related. The classroom materials are keyed to each other, to the teaching emphases, to the commentaries, and to the didactics and content sections by using "teacher talk" in non-reproducible blue print on the student pages. The commentaries refer to specific classroom pages and often a classroom page is shown reduced in size next to the discussion of the page. The commentaries relate the various teaching emphases to the statistical topic of the sub-section and to the content and didactics sections. Each teaching emphases includes a rationale and examples from the classroom materials. The concepts and skills sections contain references to the applications sections to point out the potential uses of statistics. The didactics and content sections also refer to specific classroom pages and show a classroom page reduced in size to illustrate a topic of the sections.

The resource can be used by individual teachers, teams of teachers within a school or district, and by inservice leaders for classes and workshops. Teachers can use the resource

- To learn more about statistics. The content section has many ideas and useful background information for teachers who want to review or learn concepts in statistics and probability. There are several exercises included to help teachers improve their knowledge of the subject.
- To gain insight into difficulties in learning statistics. For example, some students find it difficult to use a systematic approach in listing permutations. Difficulties such as these are discussed in the didactics sections.
- To obtain information about curriculum trends and research in mathematics education. Many of the current trends in middle school mathematics are discussed in the teaching emphases sections and keyed on the classroom pages. The didactics section contains accessible synopses of the available research related to teaching statistics and probability.
- As a source of ready-to-use activities to supplement and vary the curriculum. Worksheets can be duplicated to provide a copy for each student; activity cards can be duplicated and laminated for repeated use; transparencies

can be made from a page to form the basis of a teacher-led discussion; or a page can be the focal point of a bulletin board or display.

- As a springboard for developing activities, sequences or lessons, or curriculum. Teachers can use the classroom pages as examples which they can alter to fit the interests of their classes. From the pages of teacher ideas and from the suggestions in the commentaries and didactics papers teachers can develop other activities and sequences of lessons.
- As a source of different ways to teach a topic. A variety of teaching approaches are included in resource. Concrete models are used in many of the classroom activities and attention is given to concept formation in the commentaries and didactics sections.
- To provide students with materials according to their interests. For example, there are classroom materials for students who are interested in applied problems, students who want to be challenged, and students who need experience with concrete activities in experimental probability.
- To improve attitudes. Many activities can be used to help students have a feeling of success and accomplishment. Some students who do not succeed in arithmetic can be successful in gathering data and using the self-checking pages. The many visual and hands-on materials can be used to capture student interest.
- To increase problem solving abilities and build basic skills. Many classroom pages provide problem solving activities and the problem solving teaching emphasis contains several hints and background information.
- To gain access to other sources. The annotated bibliography describes the features of various available references for teaching and learning statistics. Many of the sources are cited in the classroom pages, and in the other sections of the resource.

The resource was made available to teachers in the United States and Canada in 1978. The reaction was varied. At one extreme teams of middle school teachers reorganized their entire curriculum by integrating the resource materials with their traditional program. At the other extreme teachers ignored the resource entirely. It was more common for teachers to phase in the resource gradually by first trying a few classroom pages, then organizing a sequence of lessons, then using one of the teaching emphases as a theme for a unit of study, and finally developing materials of their own for their students.

To some extent the resource has influenced teachers, supervisors, curriculum directors and authors of articles and textbooks just as they influenced the developers of the resource, and yet the resource is far from a panacea. It offers to the more creative, more energetic teachers splendid opportunities to enrich the learning experiences that they offer to children. In many ways the resource is too flexible especially for those teachers who only use lessons that are completely prepared for them.

While statistics is a subject in its own right, it is also a fiber that interrelates all other school topics in the social sciences, the sciences, the arts, mathematics and especially in the societal demands on people. The aspects of statistics as a truly interdisciplinary subject should be emphasized as new teaching materials are developed. Since teachers at the school level usually have the responsibility to teach several subjects, the lower grades are a natural place to begin a rich program of statistics edcuation. However, the materials that are developed should take the format of sequences of lessons that use statistics in analyzing questions in other subject areas. This connection with familiar and standard topics will make more palatable the inclusion of statistics or any other subject that may be new to the teachers.

In light of the many excellent sources that are already available [see Shulte, 1981], the materials that need to be developed now should be part of a thorough curriculum revision with a strong statistics component beginning at the school level and continuing throughout the secondary school [see, for example, Usiskin, 1979] as well as a similar integration of statistics in teacher education programs.

References

[1] Hoffer, A. (director) (1977) Statistics and Information Organization. Palo Alto, CA, Creative Publication.

[2] Shulte, A. (editor) (1981) Teaching Statistics and Probability. Reston, VA, National Council of Teachers of Mathematics.

[3] Usiskin, Z. (1979) Algebra Through Applications. Reston, VA, National Council of Teachers of Mathematics.

DEVELOPING MATERIALS FOR THE SCHOOLS COUNCIL PROJECT ON STATISTICAL EDUCATION

P Holmes, University of Sheffield

The Schools Council Project on Statistical Education (1975-1981), working in England and Wales and based at the University of Sheffield, was the first national project in the world to look specifically at the statistical education needs of all pupils in the 11-16 age range and to develop teaching materials to meet these needs. Although it was working at the same time as groups in France and in Hungary there was no connection between these projects at a time when they could have influenced each other [1].

The brief of the Schools Council Project was:

1 to assess the present situation in statistical education as regards content, level, motivation and teachers' attitudes, and to relate these to the position of statistics outside schools;

2 to survey the needs of teachers, whether they are teaching statistics as a specialist subject or working in a related field; implications for the needs of both initial and in-service training needs are to be considered;

3 to devise detailed proposals for the implementation of the teaching ideas;

4 to produce teaching materials such as notes for teachers, descriptions of experiments, work cards, sets of examples and audio-visual aids.

No curriculum development project works in a vacuum. All have to work in a particular country, taking into account the needs of society, the nature of the subject and its place in the curriculum, the background and the needs of the teachers and the needs of the pupils. In this talk I will take each of these four areas in turn and show how they affected the teaching materials developed by the Schools Council Project on Statistical Education. At the end of the talk I shall indicate the final result by giving some details from some of the units [5].

SOCIETY

Britain is a well developed society in which statistics plays a major part. There is a major Government Statistical Service whose findings are widely publicised. There are regular censuses of the population and all of the population has to fill in forms for statistical purposes at one time or another. Opinion polls are carried out frequently on many topics and many of their conclusions made public in the national press and television. People are expected to be able to understand data presented to them in popular articles, in advertising and in the financial and commerce columns of newspapers. They have to use statistics in their own places of work to make decisions. A further account of this background to the work in the project is given in [4]. Clearly the teaching material developed by the project has to reflect this national picture.

At the time the project was working there was a growing national concern about levels of numeracy amongst young people. This concern showed itself in the setting up of a national committee under the chairmanship of Dr. W.H. Cockcroft and the committee's report has recently been published [2]. In it there are a number of references to statistics. In particular it states that "such textbooks (in statistics) as are available concentrate on theory rather than practice; it seems clear that there is a need for the provision of further teaching materials which will emphasise a practical approach to the teaching of statistics" (para. 780). The Schools Council Project was concerned to produce just this type of teaching material. Another indication of the way the project and the enquiry were thinking on the same lines, and both reflecting the needs of society, is shown in para. 781. Here the committee quote approvingly the evidence submitted to it by the project that "Statistical numeracy requires a feel for numbers, an appreciation of appropriate levels of accuracy, the making of sensible estimates, a common sense approach to the use of data in supporting an argument, the awareness of the variety of interpretations of figures and a judicious understanding of widely used concepts such as means and percentages. All these are part of everyday living. "Good statistics teaching can encourage pupils to think in these ways". Again, obviously, the project was trying to develop materials that would satisfy these criteria with an emphasis on discussing accuracy of data, being involved with collecting data so as to be more aware of the problems involved and on drawing practical inferences from the results. It is shown in the concern the project has for choosing everyday examples (Sampling the Census [6], Equal Pay [7], Smoking and Health [8]), emphasising real data and the real implications of the inferences that can be drawn, with less emphasis on introducing techniques for their own sake. The project allowed and encouraged the use of calculators and, for example, provided partly drawn graphs so that more time could be spent on discussing the meaning and implications of the figures.

National needs require that most of our pupils have a general understanding of statistics. They need to know what can be done with statistics, they also need to be able to identify misleading uses of statistics. For this latter purpose the project developed a special unit on misleading data [3]. Other pupils will need to go further and develop skills in a proper use of statistics. For them it was important to devise materials that would lay the foundations of the nature of the subject - getting accurate data, allowing for variability and uncertainty, drawing inferences etc. - so that later formal work would fit into a well established schema in the pupil's mind.

Curriculum development takes place against the background of what a country expects from its education system. In England and Wales over the last few years the trend of opinion has been more towards the pragmatic, utilitarian and useful rather than the academic subject for its own sake. One of the biggest influences on the school curriculum in England and Wales are the national examinations set for 16 year olds by a few examination boards. These examinations are subject oriented, there are examinations specifically in statistics, but they are taken by only a minority of students. There is some statistics as part of many mathematics courses but, as is stated in the Cockcroft report (para. 776) "When statistics is taught within secondary mathematics courses too much emphasis is very often placed on the application of statistical techniques, rather than on discussion of the results or ordering and examining the data and on the inferences which should be drawn in the light of the context in which the data have been collected. The work can therefore become dry and technique-oriented and fail to show the power and nature of statistics". The project decided not to draw up a syllabus in statistics for one of the national examinations. This would have reached only a minority of pupils and would also have been very time consuming. It may not even have been possible within the three years for which the project was initially funded. Instead it was decided that we should try to influence courses as they currently existed, especially mathematics since this was a subject taken by nearly all pupils. The project therefore conceived of its work as a coherent series of booklets that could be used individually within different courses. The aims were to improve present content and practice, to develop what was desirable rather than what was directly assessable under examination conditions, to encourage more practical work and indirectly to change the prevailing national view of statistics particularly of mathematics teachers, mathematics examiners and those responsible for drawing up the national syllabuses.

The accepted pattern for curriculum change also affects the way teaching material is developed. In some countries the curriculum is controlled centrally; in England and Wales the final decision rests with the schools (though clearly there are regional and national pressures for conformity). This means that we had to involve schools and teachers with the development work from a very early stage. This is very good for the project team. Teachers have to be convinced that the material is useful, sensible and will work with their pupils. At the same time they make practical suggestions as to what will satisfy these criteria. By working with teachers in this way and letting them advise on, test and evaluate the trial material there is less likelihood of the material being seen as something unwanted imposed from the outside. It helps the material match up with current classroom practice. It also builds up a band of informed teachers sympathetic to the aim of the project.

THE SUBJECT AND ITS PLACE IN THE CURRICULUM

A start was made on indicating the current place of statistics in the school curriculum in the previous section. Part of the early work of the project was to investigate this position more closely. Statistics is not well established as a separate subject in the main school curriculum. On the other hand there is a widespread and growing use of statistics being made in many other subjects across the school curriculum. It also has an established part in the teaching of mathematics. This is more fully documented in the project's survey [2]. This widespread use poses several interdisciplinary problems of its own. In many cases teachers of different subjects do not know what their colleagues are doing. There are also different approaches to and requirements of statistics by teachers of different subjects. The mathematics teachers are more likely to concentrate on the mathematical aspect of the subject and seek for a mathematical understanding of the techniques. Their colleagues in other subjects may well see statistics more as a tool and not be so concerned with developing understanding. In some cases it was clear that the level of difficulty of the concept being used was not appreciated. The implications of this for the project's teaching materials were that we tried to choose topics for the booklets that could be used in different subject areas of the school curriculum and that would also encourage some interdisciplinary co-operation. Booklets developed the techniques in a context to help students, and their teachers, appreciate the why and when of using statistics. Statistics is clearly a subject that is widely applied and used in practice. The developed materials should reflect this fact.

This brings us to the impact of the nature of the subject on the materials developed. As the Cockcroft report says, quoting evidence from the project, "Statistics is not just a set of techniques, it is an attitude of mind in approaching data. In particular it acknowledges the fact of uncertainty and variability in data and data collection. It enables people to make decisions in the face of this uncertainty". All this was reflected in the project's aims.

1. Children should become aware of and appreciate the role of statistics in society. That is they should know about the many and varied fields in which statistical ideas are used, including the place of statistical thinking in other academic subjects.

2. Children should become aware of, and appreciate, the scope of statistics. That is, they should know the sort of questions that an intelligent use of statistics can answer, and understand the power and limitations of statistical thought.

These two aims can be seen as wanting children to know about the breadth of statistical applications and the depth of statistical thought. Clearly statistics introduces some very difficult ideas. These include indeterminism, uncertainty, randomness, a sample to represent a population, making estimates from samples, drawing inferences and so on. These ideas take some time to mature and occasionally run counter to the pupil's previous experiences at school. This is particularly true of the ideas of indeterminism and inference in the mathematics lessons where most of the answers are very deterministic and the emphasis has been much more on the certainty of logical deduction. Because of this the project's materials spend much time on practical work, on experiencing the problems of getting the data, of seeing not only the different outcomes that might happen but also that some are more likely than others. There is work on simulation, sampling and discussion of possible implications well before any such ideas are formalised. The practical nature of statistics encouraged the project to produce materials which taught techniques in a context and not to isolate them. This helps the pupil to decide when to use the technique and to see more clearly how it is used. Because ideas take time to mature they are introduced in an elementary way in the early units, they then occur again in later units and so the ideas are gradually reinforced. A full summary of which ideas occur in which units is given in the teachers handbook [4].

The relationship between probability and statistics also had an impact on the way the materials are being developed. At a later stage probability models, particularly probability distribution, can be seen as models for different measures in a population and hence we can quantify our inferences drawn from samples. This is far too sophisticated a concept for eleven year olds, yet probability ideas are important and take a long while to develop. Thus the early project units develop probability and statistics mainly in separate units - though there is one early unit [10] which starts to bring the two ideas together through the concepts of fairness and randomness. The early units in probability emphasise uncertainty, using games and simulations to see that decisions will sometimes be right and sometimes wrong but to get a feel that some decisions will be better than others. The early units in statistics concentrate on getting accurate data, seeing what it tells you, summarising and presenting it in different ways and drawing informal inferences from these summaries. Only in the later units do the two strands start to come together and more formal work in estimation and significance testing etc. is left to later.

THE TEACHERS

In the last resort any teaching material developed has to be accepted by, and usable by, the teachers. The initial survey carried out by the project showed that even amongst the mathematics teachers more than one-third had had to study their statistics on their own, not as part of their initial training. This proportion was even higher for teachers who had been in the profession for some time. Even for those teachers who had had some training in statistics it was usually only a small part of their background, it tended to be rather academic and did not greatly help them identify and appreciate what was suitable for their pupils. Very few had ever used statistics in a practical context. Very few indeed had had any training in the teaching of statistics, such work playing no part in by far the large majority of initial teacher training courses. There was a general feeling that statistics was a 'difficult' subject and so a lack of confidence when it came to teaching. This meant that in developing its materials the project had to think carefully how it would also help the teachers. One result was the production of teachers' notes for each of the units and a teachers handbook [4] to give background and added support.

In order to build up confidence the project team not only regularly visited the trial schools but also arranged local and national meetings for the teachers to share their problems and to become more involved with the project. Because most teachers felt their lack of competence it was not possible to persuade them to write the initial drafts, though they did give useful advice, and this writing was done by the project team. Their willingness to help with testing, evaluating and making positive comments certainly improved the quality of the material, and some of them wrote the revised version of the units prior to final editing.

Teaching styles differ from teacher to teacher and from school to school. The project worked with teachers whose natural styles varied from the stand up and talk method to the organising of completely individualised learning schemes. The teaching material had to be developed in such a way as to match up with as large a range of these styles as possible - subject to the nature of the subject allowing such a style to be sensible. The project's materials therefore include, in the pupil booklet, all the information that a pupil needs so that it is possible to use the teaching style which encourages pupils to work together with less teacher involvement. Completely individualised learning was not found appropriate to an approach which required practical experimental work, collecting data and discussing the results. Teachers used to a more formal presentation of their lessons can still work in this way but, in order to encourage greater pupil participation, there are places in the material where discussion of the results is required or there are open-ended questions where it is inappropriate for the teacher to tell the pupils "the answer". In practice it was found that even the low ability pupil had something to offer to such discussions either from practical experience or by imaginative comment.

THE PUPILS

The most important consideration of all curriculum development of teaching materials is the pupils themselves. In particular the materials must take into account the age, ability, background, present and future needs of the pupils. The Schools Council Project was developing materials for pupils of all abilities from the ages of 11 to 16. It could not assume that any earlier work had been done by the pupils on probability and statistics, though some pupils had in fact met some elementary ideas on collection and representation of data in their earlier years at school. Between the ages of 11 and 16 pupils mature greatly and their interests change and broaden. The project's material reflects this by choosing topics more from home and school in the early units where elementary ideas of probability and statistics are being developed and moves to topics of more general, social, topical and personal interest in the later units.

Because pupils of all abilities are envisaged then each unit has to include something that all can do. This might be the easier parts; it is also the sections on open questions and discussion which can be answered at many levels. Most units also incorporate extra material for the more able and also have different stopping points for those who can't work as quickly.

The present and future needs of the pupils may be for use in other subjects, in their everyday life, as people who use statistics at work or (in a minority of cases) as future specialists. The effect of these considerations on the teaching material produced has been considered above.

In this paper I have tried to show some of the many things anyone developing new teaching materials has to take into account. I have indicated how they affected the teaching material of the Schools Council Project on Statistical Education and the approach taken by the project team. I hope that this will help any of you who have to develop teaching material in your own countries to identify the areas of concern and produce relevant material for your pupils.

References

[1] Barnett, V. (1982) Teaching Statistics in Schools Throughout the World, International Statistical Institute, Voorburg.

[2] Committee of Inquiry into the Teaching of Mathematics in Schools (1982) Mathematics counts: Report of the committee of Inquiry... under the chairmanship of Dr. W.H. Cockcroft, HMSO, London.

[3] Holmes, P., Kapadia, R., Rubra, G.N. and Turner, D.E. (Ed.) (1981) Statistics in Schools 11-16: A Review, Methuen, London.

[4] Schools Council Project on Statistical Education (1980) Teaching Statistics 11-16, Foulsham, Slough.

[5] Schools Council Project on Statistical Education (1980), Statistics in Your World, (17 Pupil Booklets, 27 Teachers' Booklets) Foulsham, Slough.

[6] Schools Council Project on Statistical Education (1980) Sampling the Census, Foulsham, Slough.

[7] Schools Council Project on Statistical Education (1980) Equal Pay, Foulsham, Slough.

[8] Schools Council Project on Statistical Education (1980) Smoking and Health, Foulsham, Slough.

[9] Schools Council Project on Statistical Education (1980) Phoney Figures, Foulsham, Slough.

[10] Schools Council Project on Statistical Education (1980) Being Fair to Ernie, Foulsham, Slough.

EDA FOR 16-19's?

J Bibby, The Open University

ABSTRACT

This paper asks various questions concerning the relevance of Exploratory Data Analysis (EDA) ideas for teaching 16-19 year old's, and reports on a workshop which considered these questions.

INTRODUCTION

Four key questions can be posed which are pertinent to the question of preparing EDA materials for use at school level:

(a) are EDA ideas relevant at this level?
(b) if so, then which ideas and techniques should be taught?
(c) which data sources are sufficiently motivating for this age group?
(d) how should or could such material be prepared?

Before attempting to answer these questions it is important to be clear what EDA does and does not do.

STEMPLOTS

This simple technique provides an overview of the distribution of values in a batch. In particular, it provides an impressionistic overview of

- the general location of the batch (the "main body" of data)
- its spread, and in particular whether one batch is more, or less, spread out than another
- the location of peaks or modes
- the existence of tails, and the relative skewness or symmetry of the batch
- the existence of outliers.

However, this overview can be misleading if care is not taken to experiment with the effect of different grouping intervals.

In addition, the process of drawing the stemplot pushes the analyst's nose "close to the data", and should bring to light

- patterns related to the order of the data, e.g. auto-correlation
- any clearly aberrant observations.

In short, drawing the stemplot often provides an informal check for errors.

FIGURE 1: Measurements made by 4 individuals when attempting to bisect 12 lines by eye (Source: Kanji, 1982).

	A	B	C	D
1	3.04	3.32	3.26	3.11
2	3.16	3.23	3.14	3.10
3	3.10	3.28	3.20	3.15
4	3.01	3.35	3.18	3.16
5	3.12	3.26	3.23	3.12
6	3.07	3.20	3.27	2.98
7	3.13	3.24	3.19	3.04
8	3.07	3.32	3.25	3.04
9	3.16	3.19	3.22	3.26
10	3.18	3.26	3.19	3.20
11	3.19	3.41	3.18	3.14
12	3.27	3.03	3.22	3.13

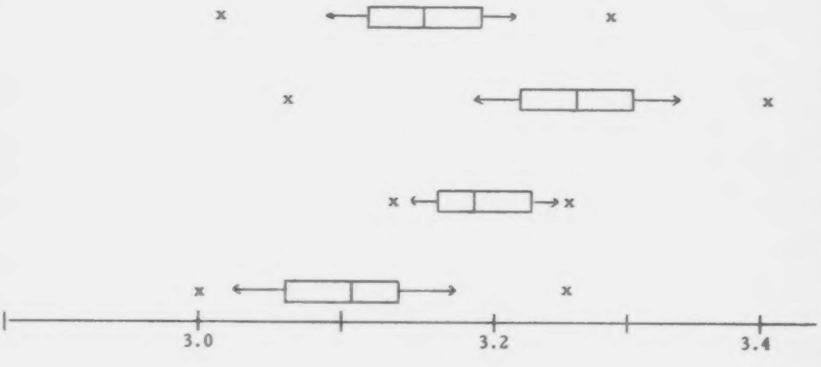


FIGURE 2: Decile boxplots based on Figure 1
(arrowheads denote deciles)

BOXPLOTS

Decile boxplots provide an indication of the distribution of data values in a batch. For example, the data in Figure 1 can be shown more graphically in Figure 2. (For further details of decile boxplots see The Open University 1983, Unit A2).

GENERAL POINTS CONCERNING EDA

There are of course many other EDA methods, e.g.

- line-fitting
- median polishing
- transforming for symmetry/linearity/additivity

All these are relatively simple, but may take time to trickle down to the school level.

The pro's and con's of EDA may be summarised as follows:

The Pro's of EDA

- EDA emphasises pattern, which is an intuitive "fuzzy" action (unlike probability)
- EDA calculations are generally simple; they use counts, not sums : hence there is no need in EDA for the intimidating panoply of sigma signs which are required by conventional methods
- EDA methods are generally applicable as they do not require a probability model; therefore they apply also where the data does not arise from a sample, where the sample design is complex, or where there has been a complete census
- EDA encourages emphasis on residuals, the DFR equation (Data = Fit + Residual), the notion of statistical analysis as an interative procedure, and the model-building loop (see An Open University Service Course, Figure 5, by John Bibby, this volume).

The Con's of EDA

However, EDA does have certain disadvantages, including the following:

- many EDA procedures are difficult to define precisely and algorithmically (e.g. deciding whether to 'squeeze' or 'stretch' a stemplot)
- terminology in the area is far from standard
- there may be problems in finding suitable teachers (it is important that statistics should be teachable even by bad teachers!)
- EDA is dangerously empiricist - it risks encouraging the notion that knowledge somehow "arises out of the data"; it downplays prior knowledge and the role of theory.

RAPPORTAGE FROM WORKSHOPS

Workshops were set up to consider questions (a) to (d) outlined in the Introduction. They are reported as follows:

There seems to be some place for EDA ideas at this level, although an immediate problem is the unfamiliarity of teaching staff with EDA ideas and techniques. There may be a need for training courses in this area. Also, certain EDA material is more suitable for younger children.

The question of suitable data sources was reckoned to be a key problem. These could be self-generated, as in the experimental task arranged by Gopal Kanji (1982) in which participants were required to bisect by eye each of a set of straight lines. Errors could be related to:-

(i) right/left handedness
(ii) sequence number on page
(iii) position on page (left/right)

Data relating to pupils' families was also reckoned to be relevant, as were large data sets such as those collected by Governmental surveys.

The preparation of EDA materials obviously had to involve classroom teachers. However, it also required some central initiative, such as could be provided by the Leverhulme Project directed by Peter Holmes at Sheffield University.

References

[1] Kanji, G. (1982) Developing materials for 16-19 year old students. (Paper prepared for ICOTS (Mineo, 5pp)).

[2] The Open University (1983) MDST242: Statistics in Society (16 vols.) Open University Press.

TOPIC 6: TRAINING OF TEACHERS IN STATISTICS

ORGANISER: PROFESSOR A P SHULTE

PREPARATIONS FOR COLLEGE TEACHING IN STATISTICS

A K Gupta, Bowling Green State University

INTRODUCTION

In recent years there has been considerable and growing interest in the use of statistics throughout many sectors of our society. The quantitative nature of this discipline has created opportunities for scientists to participate in the new technology of the twentieth century (Rao, 1973), characterized by its unique methodology, akin to what is termed "the scientific method" in the philosophy of science. This increased demand for statistical knowledge (see Amstat News, 1981) can be met adequately if the universities will produce qualified college teachers of statistics. Here we are concerned with the training of trainers. The subject is broad and many statisticians have written about the teaching of statistics and the need for reform, beginning, perhaps, with Wishart (1939 and 1948) and including Yates and Healy (1964), Rao (1972), Hogg (1972), Kempthorne (1974), Daniels (1975), Gupta (1982) among others.

Let us first look at the profile of a college teacher of statistics. A college teacher of statistics is required, with rare exceptions, to have a doctorate. He enters a college in which a pattern already exists, and in the beginning his main concern is to teach already existing elementary courses in mathematics and courses in statistics. He works in an academic atmosphere and usually works alone on his research problems. Being a specialist, he is also engaged in consulting, writing referee's reports for journals and reviewing. He is also called upon to contribute his time and energy to the work of the committees in which much important activity of an academic community is centered. Such are the demands of a college teacher of statistics. In this paper some of the basic knowledge and skills in statistics that are needed in order to prepare oneself to teach statistics effectively are discussed. Besides statistics, competencies in mathematics, probability, computational techniques, history and a cognate area are identified. The importance of participation in undergraduate teaching, statistical consulting, and continued professional development is also discussed.

REQUIRED COMPETENCIES

The prospective college teacher, as a graduate student, has the desire to learn and the ability to learn, but must also have the desire to teach and the ability to impart knowledge in a logical way. But teaching of statistics calls for expertise and hence in-depth understanding of the subject is the best preparation. The rationale here is that college teaching is a skill which can be acquired by receiving instructions. But what kind of instruction should a prospective college teacher of statistics pursue? In the light of the profile of a college teacher of statistics described above we list below the competencies which should be acquired in order to do the job effectively.

(i) Mathematics. A sophisticated training in relevant mathematics is a prerequisite to the advanced study of statistics. The competency should be gained in (a) advanced calculus (including special functions and transformations), (b) linear algebra (including vectors, matrices and eigenvalues), (c) differential equations and (d) real analysis.

(ii) Probability. Since all branches of statistics use probabilistic arguments, competency in the following topics should be acquired: (a) basic concepts (especially conditional probability), (b) stochastic processes (including time-series analysis), (c) asymptotics and (d) distribution theory.

(iii) Statistics. There are many different branches of statistics and statistics is still proliferating. The main point here is that graduate students be exposed to different points of view. Courses should emphasize modern statistical theory and its applications and include (a) statistical inference (estimation and testing), (b) nonparametric analysis, (c) Bayesian analysis, (d) sequential analysis, (e) decision theory, (f) multivariate analysis and linear models (including regression), (g) data analysis, (h) sampling techniques and design of experiments.

(iv) Computing Techniques. Computers are used as an aid in teaching statistics, analyzing data and simplifying calculations. Hence knowledge of at least one programming language (e.g. Fortran, Basic, APL) and familiarity with most useful statistical packages (e.g. BMDP, SAS, SPSS, MINITAB, etc.) is imperative. Courses should include (a) programming, (b) data processing and (c) simulation techniques.

(v) Cognate Studies. Statistics is a dynamic subject which pervades so many fields including biological, medical, physical and social sciences. In fact a great many problems that statisticians solve come from these fields. Competency, maybe as a minor, in at least one such allied field should be acquired.

(vi) History. The history of statistics is a fascinating one and every teacher should know it. Knowledge of history gives the teacher confidence and makes teaching lively. Every statistician should be familiar with the works of Galton, K. Pearson, Fisher, E. Pearson, Neyman and other pioneers.

PARTICIPATION IN UNDERGRADUATE TEACHING AND STATISTICAL CONSULTING

So far we have been concerned with the coursework training of prospective college teachers of statistics. However, teaching also belongs to the activities people learn by doing. It has been suggested (MAA, 1979) that teaching assistantship should be teaching apprenticeship during which prospective college teachers (then graduate students) get training in the skills that will make them effective in the classroom. Next to the competencies to be acquired through coursework, participation in the supervised undergraduate teaching is the best preparation for a college teacher.

Another mode of preparing by doing is the participation in statistical consulting. This is usually achieved by working in a statistics laboratory or consulting center at a unviersity. By working on real problems, a student learns the art of asking questions, reads relevant material on the topic and gains experience in the art of communication so important for a statistician and college teacher.

IN-SERVICE TRAINING

A college teacher of statistics needs to keep up with the changes in the subject matter, advances in instructional technology and new societal demands. He needs skills of curriculum planning. Competency in these areas can be achieved through in-service training and continued professional development programs. In this context the concept of useful half-life is pertinent. For example, an engineer is thought of as losing half of his effectiveness each 8 years. Hence universities and professional societies should create the mechanism for training and refresher programs for college teachers of statistics. The American Statisical Association and the American Mathematical Society do provide short courses in relevant new topics in statistics (and mathematics), some of which are available in the form of films and tapes. It will be very useful to establish some sort of clearing house to publicize the availability of such audio and video materials so that college teachers of statistics throughout the world could benefit.

CONCLUDING REMARKS

The universities in the United States are doing an adequate job of preparing college teachers in statistics. However, improvements are necessary. There can hardly be a single international approach to the education of college teachers in statistics, though one should expect common international standards of excellence in statistical education. Effective college teaching may be independent of the subject but a regional approach is very important in preparing for college teaching in statistics due to factors such as: resources, language, etc. However, international collaboration among nations in preparing college teachers in statistics could be of great value. The International Latin American Training Center in Statistics involving 15 countries is an example of such a collaboration. In this regard an International Directory of Academic Statisticians is proposed to be made available. Such a directory should include, in addition to their specialty areas and research interests, information about their linguistic capabilities. Such a reference source, I hope, will be useful in drawing from the world pool of trained statisticians, to help prepare college teachers in statistics for regions where they may be needed.

References

[1] Daniels, H.E. (1975) Statistics in universities - a personal view (with discussion). J.Roy.Statist.Soc., Series A, 138, 1-17.

[2] Gupta, A.K. (1982) Random thoughts on developing statistics in the thrid third world. Trabajos de Estadistica.

[3] Hogg, R.V. (1972) On statistical education. American Statistician, 26, 8-26.

[4] Kempthorne, O. (1974) The teaching of statistics at the M.S. level. An essay prepared for the inaugural ceremonies for the Center of Statistics and Computation, Chapingo, Mexico.

[5] Rao, C.R. (1972) A multidisciplinary approach for teaching statistics and probability. Chapter 13, pp 251-272 in The Teaching of Probability and Statistics. Edited by L. Råde.

[6] Rao, C.R. (1973) Linear Statistical Inference and its Applications. Wiley, New York.

[7] "Supply and Demand for Statisticians". (1981) Amstat News, April, No.74, 5.

[8] Training Programs for Teaching Assistants in Mathematics, (1979) Mathematical Association of America.

[9] Wishart, J. (1939) some aspects of the teaching of statistics. J.Roy.Statist.Soc., Series A, 102, 532-564.

[10] Wishart, J. (1948) The teaching of statistics. J.Roy.Statist.Soc., Series A, 109, 532-564.

[11] Yates, F. and Healy, M.J.R. (1964) How should we reform the teaching of statistics? J.Roy.Statist.Soc., Series A, 127, 199-210.

TEACHING DESIGN OF EXPERIMENTS IN INDIAN UNIVERSITIES

V Bhagwandas, University of Indore

A GLIMPSE OF INDIA

India is a vast country, the Indian population being 15.53% of the World's population (reference year 1980-81) and the second largest country after China (21.72%). This country is divided into 22 states and 9 union territories inhabited by 339.9 million males and 318.2 million females with percent decadal growth rate 1971-81 as 24.64%. Increase in population density from 177 persons per square kilometer in 1971 to 221 persons per square kilometer in 1981.

Literacy in the Census of India is defined as follows:

The person who can read, write and understand a language is defined as literate whereas the person who can only read a language but cannot write it is supposed to be illiterate. Children below the age of five years are deemed to be illiterate. On the basis of working paper I (1981) Census of India (which includes children below five for 1981 data; due to absence of this information separately till the publication of this paper), the literacy rate is as follows:

	Persons	Male	Female	Male/Female Ratio
1961	24.02	34.44	12.95	941
1971	29.45	39.45	18.69	930
1981	36.17	46.74	24.88	935

*Little more cross section for 1961 and 1971 (1981 data is not available) for the age-group 15-35 i.e. adult group has the following data on literacy rate.

Literacy position of the population in the age group 15-35 by sex and Urban/Rural areas in 1961 and 1971 as follows:

		% of Illiterate			% of Literate		
Year	Area	Persons	Male	Female	Persons	Male	Female
1961	Total	67.60	53.17	82.45	32.40	46.83	17.55
	Urban	39.28	26.98	54.41	60.72	73.02	45.59
	Rural	74.61	60.37	88.63	25.39	39.63	11.37
1971	Total	57.97	43.29	73.24	42.03	56.71	26.76
	Urban	31.40	21.47	43.33	68.60	78.53	56.67
	Rural	65.75	50.27	81.20	34.25	49.73	18.80

---

* Source: Literacy Statistics at a Glance, Directorate of Adult Education

The definition of urban area adopted after the 1951 Census is as follows:

(1) All places with a municipality, corporation, Cantonment board or notified town area committee etc.

(2) All other places which satisfy the following criteria:

   i) a minimum population of 5,000
   ii) at least 75% of the male working population engaged in non-agricultural pursuits
   iii) a density of population of at least 400 persons per square kilometer

We give Number of Institutions for 1978-79 and 1979-80 and enrolment by stages/ classes for the year 1978-79 as Table 1 and Table 2 as given by Planning, Monitoring and Statistics Division, Ministry of Education and Culture, Government of India "Studies in Educational Statistics No.1 - 1981".

Under the Constitution of India, education is a primary responsibility of the state and states have compulsory education for children between the ages 6 to 11 and that explains the large primary schools in the country. High School/Higher Secondary and Intermediate or college education is imparted both by private as well as government schools and colleges. Unfortunately, the quality of education imparted at government schools/colleges is poorer as compared to private and a commoner cannot afford higher fees at private institutions. Just to give an idea of pass percentage in High School and Higher Secondary School, after looking into the results of various Boards of Secondary Education of States in 1977 we give the following table:

Table 1

NUMBER OF INSTITUTIONS

Type of Institutions	1978-79	1979-80
(a) Pre-Primary Schools	9,172	10,426
(b) Primary Schools	474,993	478,249
(c) Middle Schools	110,902	114,720
(d) High Schools	36,190	37,419
(e) Higher Secondary Schools (Old Pattern)	3,121	3,127
(f) Higher Secondary Schools (New Pattern)	4,531	5,044
(g) Inter/Junior Colleges	3,017	3,197
(h) Arts, Science & Commerce Colleges	3,261	3,336
(i) Universities	106	108

Table 2

ENROLMENT BY STAGES/CLASSES (1978-79)

(INDIA)

	Boys	Girls	Total
Ph.D./D.Sc./D.Phil.	16,376	5,817	22,193
M.A.	106,417	64,350	170,767
M.Sc.	48,529	22,458	70,987
M.Com.	31,719	2,001	33,720
B.A./B.A.(Hons)	626,231	360,494	986,725
B.Sc./B.Sc.(Hons)	381,346	136,069	517,415
B.Com./B.Com.(Hons)	412,801	58,538	471,339
B.E./B.Sc.(Eng)	94,501	3,246	97,747
B.Ed./B.T.	43,761	32,199	75,960
M.B.B.S.	53,617	16,717	70,334
Inter Arts	190,617	65,055	255,672
Inter Science	158,455	35,023	193,478
Inter Commerce	39,124	2,201	41,325
Pre-Degree/Pre-University (two-year course)	186,739	100,365	287,104
Pre-University (one-year course)	57,143	22,215	79,358
Higher Secondary (New Scheme XI and XII Classes)	807,670	386,284	1,193,951
Higher Secondary (Old Scheme IX, X and XI Classes)	945,313	303,765	1,249,078
High School (IX and X Classes)	4,854,358	2,066,915	6,921,273
Middle (VI-VIII Classes)	12,548,337	6,152,894	18,701,231
Primary (I-V Classes)	43,662,003	27,278,383	70,940,386
Pre-Primary Stage	488,961	415,170	904,131
Teacher Training Schools	59,411	45,713	105,124
Polytechnics	116,907	10,845	127,752
Technical and Industrial Schools	163,185	37,701	200,886

	Number Appeared			Number Passed			Percentage		
	Boys	Girls	Total	Boys	Girls	Total	Boys	Girls	Total
High School (Annual)	2,549,794	983,064	3,532,858	1,141,566	427,898	1,569,464	44.8	43.5	44.4
High School (Suppl.)	555,722	146,548	702,270	163,663	39,352	203,015	29.5	26.9	28.9
Higher Secondary (Annual)	416,218	144,775	560,993	196,478	91,150	287,626	47.2	63.0	51.3
Higher Secondary (Suppl.)	44,697	21,946	66,647	17,792	9,342	27,134	39.8	42.6	40.7
Intermediate Jr./Pre-university (Annual)	387,041	106,987	494,028	229,248	68,793	298,041	59.2	64.3	60.3
Intermediate Jr./Pre-University (Suppl.)	53,232	6,400	59,632	27,518	3,003	30,521	51.7	46.9	51.2

The pass percentage rises up as we go from high school level to university level due to students feeling more responsible; the education is costlier and not many, therefore, can afford higher education.

UNIVERSITIES IN INDIA

The first universities to be established were Calcutta, Bombay and Madras in the year 1857, the next Punjab University in 1882 and then Allahabad in 1887. For quite some time these five universities constituted the university system in India,and there were 19 universities before independence in 1947. By 1960, 45 universities were established and by 1970 the total number had risen to 83; today the number is 137, out of which 93 are traditional universities, 21 are agricultural universities and 10 belong to the category of the institutions of national importance and 12 are known as deemed to be universities,and one not covered was established in 1981, named as Jagannath Sanskrit Vishvavidyalaya.

In India, colleges were established first and universities came later. When the universities came to be established the colleges ceased to be independent units and became affiliated units of the universities. In fact the affiliating system is to be found only in India. On the recommendation of the Sadler Commission some of the unitary and residential universities, like Lucknow, Dacca, Waltair and Delhi were established but just after independence and due to a tremendous surge in enrolment most of the expansion took place along the lines of the establishment of a large number of colleges/universities which were affiliating in character and became the foci of growth and their number rose from 14 in 1947 to 74 as of today.

In recent years, research facilities have grown a great deal in range and variety and with new universities, departments of postgraduate education and research are increasing and new postgraduate colleges too try to equip themselves with research facilities. There had been a reduction in the funds available with the U.G.C. for research but other agencies like CSIR, ICAR, ICMR, Department of Atomic Energy, Department of Science and Technology, etc. do support researches in Science and Technology. But unfortunately, in India there is no rationale behind distribution of these funds and some universities and national institutes are favoured much more as compared to others.

Out of 94 traditional universities if we exclude 10 universities like Sanskrit Universities, Technology Universities, Ayurved Universities, then 84 can be classified according to type as follows:

S.No.	Type	Number
1.	Teaching and Affiliating	42
2.	Affiliating Teaching and Residential	10
3.	Affiliating cum Teaching	9
4.	Affiliating	6
5.	Teaching and Residential	5
6.	Unitary, Teaching and Residential	4
7.	Unitary and Teaching	2
8.	Federal	2
9.	Teaching	1
10.	Unitary and Residential	1
11.	Teaching cum Residential	1
12.	Unitary and Affiliating	1
	Total:	84

STATISTICS IN UNIVERSITIES

(a) 13 out of 21 Agricultural Universities have Statistics departments, 3 being departments of Agricultural Statistics. The largest of these departments is the Department of Agricultural Economics and Statistics in Punjab rao Krishi Vishwavidyalaya, Akola, having 2 Professors, 4 Readers and 17 Assistant Professors; the second largest is that of Punjab Agricultural University.

(b) 10 institutions of National importance include 5 Institutes of Technology at Bombay, Delhi, Kanpur, Kharagpur and Madras and only I.I.T. at Bombay and Kanpur have regular postgraduate courses in Statistics. One Institute, Dakshina Bharat Hindi Prachar Sabha is for the 'Hindi' language and 3 Medical Institutes, out of which the All India Institute of Medical Sciences at Delhi and the Postgraduate Institute of Medical Education and Research at Chandigarh take the services of Bio-Statisticians. Tenth is the internationally known institute viz. Indian Statistical Institute, meant for Statistics and Statisticians.

(c) Out of 12 institutions deemed to be universities the Indian Agricultural Research Institute at Delhi is the only institution having a large department of Agricultural Statistics doing useful work in teaching and research in Statistics.

(d) The following table gives the number of universities imparting postgraduate (at Master's level) instruction in Statistics.

S.No.	Type	Number of Universities having P.G. in Statistics
1.	Teaching and Affiliating	21
2.	Affiliating Teaching and Residential	7
3.	Affiliating cum Teaching	6
4.	Affiliating	3
5.	Teaching and Residential	3
6.	Unitary Teaching and Residential	4*
7.	Unitary and Teaching	Nil
8.	Federal	1
9.	Teaching	Nil
10.	Unitary and Residential	Nil
11.	Teaching cum Residential	1
12.	Unitary and Affiliating	Nil
	Total:	46

* All the four universities which are unitary, teaching and residential have a postgraduate department of teaching in Statistics. Half of the universities in India have yet to open doors for master's level instructions in Statistics.

TEACHING AND RESEARCH IN DESIGN OF EXPERIMENTS IN INDIAN UNIVERSITIES

The importance of Statistical Designs in Statistics was established by Federer & Federer [1] in 1973, while making a case for a separate research journal in Design of Experiments and allied branches. Out of the articles published in 18 journals between 1968 and 1971, 15.6% of articles and 15.2% of page coverage was only in the area of Experimental Design, Treatment Design Measuring and Measurement, Sampling Survey Design, Sequential Design, Modelling Design and others. Now if we drop the topics like Sample Survey Design and Sequential Design, we can see that percentage still does not fall below ten.

The main centres in India for research in Design of Experiments have been the Indian Statistical Institute Calcutta, Calcutta University, Bombay University and the Indian Agricultural Research Institute, Delhi and the main sources of inspiration have been people like Professor C.R. Rao in Calcutta, Professor S.S. Shrikhande and Professor M.C. Chakraborty in Bombay and Professor M.N. Das in I.A.R.I., New Delhi.

The questionnaire was sent to all the 46 universities' departments and institutes to find out the status of these departments with respect to teaching and research in Design of Experiments. Only 25 departments and institutes responded - the other 21 departments did not feel it necessary to respond. The findings are as follows:

(1) Most of the universities have Design of Experiments as a compulsory subject forming half the paper along with Multivariate Analysis, Matrix Theory or Sampling Theory except in Madras, where Design of Experiment is also offered at M.Phil. level and at optional level. Four universities, namely Calcutta,

Punjabi University Patiala, Gauhati and Indore offer a compulsory as well as an optional paper in Designs. At optional level, one treats the subject at advanced level. Construction of Designs is taught in nine universities.

(2) Topics covered mostly in these courses on design include: Theory of Linear Estimation, Gauss Markoff's Theorem, Linear Models, General Theory of Analysis of Experimental Designs, Randomised Block, Latin Squares, Graeco-Latin Squares, Cross-over Designs, Incomplete Block Designs; Balanced and Partially Balanced, Youden Square, Split Plot Experiments, Symmetrical and Unsymmetrical Factorial Experiments, Missing Plot Techniques and Response Surface Analysis.

On the construction side, the knowledge of EG(n,s), PG(n,s) complete set of m.o.l.s and their connection with PG(2,s) EG(2,s). Orthogonal Arrays, Hadamard Matrices, Tactical Configurations, Duals of Incomplete Block Designs, the application of Graph Theory in Construction of Designs, Construction of various GD designs, Triangular Designs, L_r association Scheme etc. is usually covered in a good course.

(3) Books followed in order of preference are given below. The numbers in brackets indicate the number of departments using the books.

(1) Oscar Kempthorne	"Design & Analysis of Experiments" (21)
(2) Cochran & Cox	"Experimental Designs" (16)
(3) Chakraborty, M.C.	"Mathematics of Designs and Analysis of Experiments (13)
(4) Raghavrao, D.	"Construction & Combinatorial Problems in Design of Experiments" (9)
(5) Das, M.N. & Giri, N.C.	"Experimental Designs" (8)
(6) Federer, W.T.	"Experimental Designs" (7)
(7) John, P.W.M.	"Statistical Design and Analysis of Experiments" (6)
(8) Mann, H.B.	"Analysis & Design of Experiments" (4)
(9) Scheffe, H.	"Analysis of Variance" (3)
(10) Rao, C.R.	"Linear Statistical Inference" (2)
(11) Jambunathan	"Linear Estimation" (2)
(12) Cox, D.R.	"Planning of Experiments" (2)
(13) Fedorov	"Theory of Optimal Experiments" Academic Press
(14) Fisher	"The Design of Experiments"
(15) Pearce, S.C.	"Field Experimentation with Fruit Trees and Other Perennial Plants"
(16) Montgomery, D.C.	"Design & Analysis of Experiments"
(17) Anderson, R.L. & Bancroft, T.A.	"Statistical Theory in Research"
(18) Panse, V.G. & Sukhatme, P.V.	"Statistical Method for Agricultural Workers"

(19) Ogawa, J.	"Statistical Theory of the Analysis of Experimental Designs"
(20) Ryser, H.J.	"Combinatorial Mathematics"
(21) Hall, Marshall	"Combinatorial Theory"
(22) Bose, R.C., Clatworthy, W.H. & Shrikhande, S.S.	"Tables of PBIBD with two associate classes"

The book numbers (20) and (21) are recommended in Nagpur University for half paper on Construction of Design due to the fact that Professor Shrikhande works in Nagpur as Honorary Professor after retirement from Bombay.

(4) Research publications in Design of Experiments during the last five years by various university departments:

1. Indore Univeristy (13) one thesis submitted
2. Madras Univeristy (10) and 3 technical reports
3. Calcutta University (9)
4. Punjabi University (9) two theses in Design
5. University of Rajasthan (8) two theses in Design
6. Lucknow University (6)
7. Sardar Patel University (6)
8. Annamalai University (5)
9. Utkal University (2)
10. Osmania University (1)
11. Bombay University (1)

Total (70) research papers + 3 technical reports.

(5) Research Project work completed by departments in Design of Experiment:

Calcutta University

(i) On PBIB association scheme
(ii) On response surface designs
(iii) On asymmetrical factorial
(iv) Balanced factorials
(v) Estimation of variance components

Rajasthan University

(i) Construction and analysis of Multi-dimensional designs.

Lucknow University

(i) Optimality of designs in Regression Analysis

Punjabi University

(i) Construction of incomplete block designs and of designs for Bio-assays
(ii) PBIB designs, their uses in genetical experiments and their construction
(iii) Incomplete block designs and their construction

Bombay University

(i) Generalised group divisible designs and their properties and generalised Kronecker Product of Designs

Indore University

(i) Construction of Bose Block Designs
(ii) Construction of Semi-regular and Regular GD Designs

(6) Some of the suggestions to improve the existing courses and research facilities are as follows:

(i) Distinguished scholars (!) should be invited for talks (usually people forget about the talk after delivery, if not basically interested in the topic of the talk).

(ii) Encouragement must be provided by the government for writing monographs on Design of Experiments.

(iii) In India, we have the tradition of a Summer-Institute organised for four to six weeks on an all-India/regional basis on specialised topics in various subjects and it is suggested that a Summer-Institute be organised in Design of Experiments at all-India level.

(iv) One of the suggestions includes separate books for analysis and construction respectively on design of experiments.

(a) It is true that the only book available on construction of design is due to Raghav Rao and a good book supplementing the content of Raghav Rao is welcome. Das and Giri's book is good for its general use but a book on Analysis with rigorous treatment is also needed in the subject of design of experiments.

(b) Most of the universities in India are very poorly equipped with respect to libraries. Even if some journals are subscribed, it is no wonder if some of its issues are missing due to non-subscriptions (yearly), the reason being irregular and insufficient flow of grant from supporting agencies like the University Grants Commission. In fact, at least one university in a State must have all issues of the most important journals of the subject. Various universities of the State may be allowed different subjects in order to avoid any confrontation.

(c) The interaction between various academicians in the same area of the subject is minimum at present in India and any effort like regional seminars/discussions etc. including participation in a Summer Institute is a necessity.

(d) Education, Agricultural Scientists, Psychologists etc. are consumers of the products of design of experiments and interaction with them is going to help in evolving new methods in design which are lacking in the country. Consumers listed above use available techniques instead of asking a design of experiment person to modify the technique to suit them. In fact in the process, sometimes they tailor out even a good piece of data.

Reference

[1] Federer, W.T. and Federer, A.S. (1973) "A Study of Statistical Design Publications from 1968 through 1971". The American Statistician, Vol.27, pp.160-163.

CLINICAL APPROACH TO TRAINING STATISTICS TEACHERS

J Rustagi, The Ohio State University

INTRODUCTION

Statistics is a relatively new field of academic and scientific endeavour. With the advent of new technology, especially the computer, the art of data analysis and statistical computing has attained a new perspective. With the emerging field of computer graphics, statistical sciences is bound to receive new and innovative techniques for solutions to many problems. In spite of substantial theoretical and experimental developments in statistics, the teaching of statistics has not received adequate attention by the profession. It is only recently that attempts have been made by the national and international statistical associations to address the issues involved. The latest such effort was made at The Ohio State University in organizing a conference in Teaching of Statistics and Statistical Consulting in 1980. The proceedings of this conference were published under the editorship of Rustagi and Wolfe (1982).

One of the important issues not addressed at that time was the training of teachers of statistics for schools and colleges. In this paper, we propose to provide a framework to enrich the training of statistics teachers at the college level. Recently the need of such training has been felt especially in the United States, after the Mathematical Association for America proposed courses in Statistics at the undergraduate level through its Committee on Undergraduate Programs in Mathematics. Such courses are being taught often by persons not really trained in the statistical sciences. We propose a "clinical" approach for training statistics teachers so as to provide "bedside" teaching of statistics, so often used in medical sciences. Such a training is especially suited to persons who have an understanding of the mathematical basis of the techniques and need to develop an appreciation of the nature of the statistical phenomenon.

STATISTICS AS AN IMPORTANT DISCIPLINE IN SCHOOLS AND COLLEGES

The applications of statistical techniques are made in a large number of scientific disciplines. The problems of basic design of experimentation, analysis of experimental results, acceptance or rejection of scientific hypotheses are extremely important to scientists. In the non-experimental sciences such as economics, the numerical information is utilized in developing models. As an important branch of knowledge, statistics is being required in the curricula of many undergraduate and graduate degree programs. In high schools an understanding of probability and statistics is required of the students and statistics courses have been introduced at high school levels in many countries. Several mathematical techniques can be motivated through statistics. Suggestions to teach elementary mathematical concepts through statistics have been made by several mathematics teachers (for reference see Anslander et al. (1971)).

At present a usual unit in high school mathematical curriculum in American high schools includes concepts in probability and statistics. In some school systems separate courses in statistics have been introduced. The television course on probability and statistics in the early sixties by F. Mosteller and further developments in the high school curriculum (Mosteller, 1971), popularized the introduction of statistics in high schools in the United States and other countries.

The requirement that students take at least one course in statistics with prerequisites of calculus has been introduced in many undergraduate business curricula. A large number of universities and colleges provide an introductory course to various non-mathematical majors with only the minimum prerequisites. Such service courses are increasing at a fast pace and teachers without substantial training in statistics are being called upon to teach them. The Committee of Undergraduate Programs in Mathematics of the Mathematical Association of America (1972) has developed course contents for statistics courses and this has created a demand for statistics teachers in undergraduate colleges. Very few universities in the United States have independent Departments of Statistics and hence this responsibility falls on teachers of mathematics, psychology, economics etc. who have a side interest in statistics. In such cases, not only is it important to provide material for teaching but also to provide techniques for teaching.

Undergraduate programs in biological, social, behavioural and engineering sciences require statistics courses as prerequisites to their courses in their own disciplines. Increasing demand is being made for courses in statistics as a result of the introduction of interdisciplinary programs such as majors in the mathematical sciences which include statistics, mathematics and computer science. There is a growing demand for statistics teachers in schools and colleges at every level. A recent survey at The Ohio State University disclosed that more than two hundred faculty members were teaching, consulting or applying statistics as a major activity of their academic endeavour even though the Department of Statistics has only a faculty of fifteen (Report of Self Study Committee, 1981).

LACK OF TRAINED TEACHERS

In the recent past there has been an increase in the creation of new Departments of Statistics programs. Several new programs in American universities have been introduced for training master and doctoral level statisticians. Many undergraduate programs have teachers without formal training and practical experience in statistics. Such teachers provide, at times, only a formal treatment of the subject. Many times, the required courses are taught to the "unwilling" student by teachers having no enthusiasm and no experience with real application. They hardly transmit any stimulation and motivation, thus depriving a large majority of students of the real potentials of statistical sciences. Many potential scientists, educators and future leaders do not receive rich heritage of the exciting field of statistics. Lack of proper training in statistics at the college level not only affects the field in which it has direct application but prohibits the students from being exposed to one of the most important modern sciences.

To perform the task of teaching statistical concepts effectively, a basic understanding of statistical philosophy and practice is needed. The competent teacher should be able to provide the feeling of the real world problems and the application of statistics to these problems. The statistics teachers should be trained to impart this feeling to their pupils in an effective manner and with enthusiasm.

CLINICAL APPROACH

In order to provide statistics teachers with the subject matter of statistics, basic concepts in the theory of probability, statistical inference, design of experiments, etc. should be conveyed. Facility of utilizing well known statistical packages for numerical computations should be developed. Understanding of the problems of data analysis and development of statistical theory for this purpose can be made through the problem-oriented approach. This approach to training statistics teachers requires that the individual be associated with an actual problem in data analysis. The enthusiasm and insight of the experimenter can be easily transmitted to the trainee if he/she is associated in a meaningful way with the experiment. The purpose of this association is that the student is guided towards the appreciation of all aspects of experimentation, development of model, formulation of hypotheses, data collection, data analysis and the final disposition of the problem.

Another way of providing such a training is through internship programs in some industrial or business organization where the student is exposed to real problems. Practical difficulty of arranging such internships and the limited contributions of such assignments to the training of statistics teachers, has made the method of little use. However, the flow of practical problems in a university setting is quite large and can provide sufficiently different problems of experimentation to give a broad survey of statistical concepts. In our experience at The Ohio State University, Statistics Laboratory, the range of problems in one quarter of 10 weeks, may see a problem in Home Economics, Anthropology, Educational Research, Engineering, Agriculture, Medicine, Music, Literature and so on. Many of these problems provide the student with a wide variety of applications of statistics to various fields.

The experimenter discusses the problem with a statistical consultant in the presence of the student. The problem is analyzed and statistical questions are formulated. The complete process of numerical evaluation of the results and final disposition of the problem is conducted with the experimenter in the presence of the student. A written report is prepared often and is discussed by all the parties.

The importance of having a statistical laboratory cannot be overemphasized. The laboratory provides a continuous flow of actual problems and gives the student teacher the experience needed. The techniques of teaching statistical consulting are also developed in a similar fashion. A series of papers on this matter is given in Rustagi and Wolfe (1982).

References

[1] Anslander, L., Avenoso, F.J., Dyer, E., Chiefetz, P., Gewirtz, A., Hunte, B.E., Quintas, L.V. and Rauch, H.E. (1971) Mathematics Through Statistics, Mathematical Dimensions, Ill., Chicago, Ill.

[2] Mosteller, F. (1971) The Joint American Statistical Association - National Council of Teachers of Mathematics Committee on the Curriculum in Statistics and Probability, Review of the International Statistical Institute, 39, 340-342.

[3] Panel on Statistics, The Introductory Statistics Courses, Committee on Undergraduate Program in Mathematics, Mathematical Association of America, Berkeley, California, 1972.

[4] Polya, George (1963) On learning, teaching, and learning teaching. The American Mathematical Monthly, 70: 6, 605.

[5] Report of the Self Study Committee, Department of Statistics (1981) The Ohio State University, Columbus, Ohio.

[6] Rustagi, J.S. and Wolfe, Douglas A. (Editors) (1982) Teaching of Statistics and Statistical Consulting, Academic Press, New York.

[7] Situmeang, R.P. (1976) Motivational supplements to lecture discussion in introductory college statistics courses, Ph.D. dissertation, The Ohio State University, Columbus, Ohio.

[8] Tanur, Judith et al. (Editors) (1972) Statistics: A Guide to the Unknown Holden-Day Incorporated, San Francisco,

DISCUSSION OF TEACHER-TRAINING PAPERS

A Morley, Trent Polytechnic

The papers from Professor Gupta and Professor Rustagi are concerned with the training of teachers of statistics at college level. They are particularly concerned with the kind of statistical training experienced by the prospective college teacher, especially in applied statistics, and this is clearly very important. However, they make limited comment on more direct pedagogic training. It may be useful for me to comment on this aspect of training for teachers at both college and school level.

Professor Gupta envisages the new college teacher taking part in established undergraduate courses with a set text. Typically in the U.K. he will be faced with much more responsibility for the presentation; often just given the syllabus. For example, there are no suitable texts for many services courses.

To help with taking this responsibility we have recently instituted required training for new teachers in state institutions for higher and further education, organised on a regional basis. In the first year the newcomer has a reduced teaching load, a three-day orientation course before teaching begins, tutorial supervision by an experienced member of his own department which includes observation of lectures, and a two-week course at the end of the year. In the second year he will attend for one day a week throughout the year a course organised by the Staff Development Officer.

As a supervisor the kind of things with which I am trying to give assistance are:

(i) *Conceptual and task analysis of a topic.*
Student background-levels of symbolism and justification to be used.

(ii) *The shaping and presentation of a lecture; posing and answering a problem - key questions - aids.*

(iii) *Planning of workshop and seminar activities.*

(iv) *The overall sequencing of a syllabus. Time to be spent on various topics. The use of texts, handouts, tables, calculators and computers.*

(v) *Setting and marking of coursework, tests and end of year examinations.*

At school level in the U.K. we have to distinguish sharply between training teachers for so-called 'A' level courses in statistics for a selective group (approximately top 15%) of 16-18 year olds and for teaching elementary probability and statistics within integrated mathematics courses for 11-16 year olds.

The number of 16-18 year old pupils taking statistics has grown enormously in recent years. The courses are ambitious, externally examined nationally and assume a concurrent study of the calculus. Those training to teach at this level will mainly be mathematics graduates with a variable knowledge of statistics, doing a one-year course of teacher-training (including 12 out of 36 weeks on teaching practice in schools). There is relatively little time to spend specifically on the teaching of statistics, typically 6-10 hours. The tutor is limited to considering some of the texts available (e.g. Hodge and Seed), the syllabus and type of examination of the different examination boards, including those where projects have a part and study of some key issues - I look at hypothesis-testing, the use of microcomputer graphics, and the Statlab data-bank. There is a great need for in-service courses for this group of teachers.

A similar small amount of time is available for considering statistics for 11-16 year olds and we have to rely on transfer from general experience in the teaching of mathematics. I focus on a few units from the 'Statistics in Your World' series looking at the notion of key questions, identifying skills and concepts, how to keep up-to-date with the data and the teachers' book which discusses the justification for statistics in the curriculum at this level.

TOPIC 7: USE OF CALCULATORS AND COMPUTERS IN TEACHING

ORGANISER: PROFESSOR L RÅDE

THE CASE FOR USING COMPUTERS IN TEACHING STATISTICS

B L Joiner, University of Wisconsin

ABSTRACT

The teaching of basic statistics courses can be greatly enhanced with the use of widely available statistical software. Pocket calculators can relieve some of the computational drudgery of statistics but are of little help with data plotting, one of the key steps in the statistical analysis process. By using computers in basic statistics courses, we can present students with real data sets and problems, and teach them how to approach and analyse data. In this paper, the basic steps in the statistical analysis process are listed, and the important role that computers can play in some of these steps is emphasised.

INTRODUCTION

What can be conveniently calculated and plotted has always played a major role in how statistics was done and in how it was taught. Until recently the best tools available to practitioners and students were mechanical calculators that could only add, subtract, multiply, divide and perhaps take a square root. These were used by many to supplement hand calculations.

In the past decade two important developments have offered new possibilities. These are pocket calculators and widely available easy-to-use computer software. Relatively inexpensive pocket calculators can now do higher level calculations such as compute standard deviations, fit straight lines and even certain types of non-linear functions, compute percent points and integrals of statistical distributions and simulate random data. Using these effectively, as is done, for example, by Koopmans (1981), can relieve much of the computational drudgery traditionally associated with basic statistics courses. However, a key weakness in these devices is their inability to speed up substantially the plotting of data.

Tukey (e.g., 1977) has introduced new graphical displays, many of which can be done by hand relatively easily. Combining these new displays with the use of pocket calculators does open exciting new possibilities for low-budget teaching, as Koopmans (1981) has demonstrated.

In the places where students can have access to computers, however, an even better approach is possible. This approach adds on easy-to-use computer languages such as Minitab (Ryan, Joiner, Ryan, 1976, 1982) to the tools traditionally available. The cost of this approach is decreasing rapidly and, in the United States, may now be as low as $.25 per half-hour session. If a student had 20 half-hour sessions in a course, this would amount to about $5 per student, or roughly one-third to one-fourth the cost of a text book.

THE STATISTICAL ANALYSIS PROCESS

"I know how to do a t-test, I just don't know when to do one."

If our teaching is to do any good, our students must learn to judge when various statistical procedures are appropriate, what assumptions are critical, how to check for violation of assumptions, how to judge whether calculations have been done at least approximately correctly and how to interpret results. In other words, they must learn the basics of how to do statistics, not just do calculations.

The statistical analysis process has many steps, most of which need to be done in loops several times in any analysis. These steps include the following:

(1) Understanding something about the context of the data
(2) Understanding the goals of the analysis
(3) Checking the data values for reasonableness
(4) Assessing whether or not the data might shed light on the stated goals of the analysis
(5) Understanding something about the sources of variation or uncertainty in the data
(6) Making a variety of plots that increase understanding of the data
(7) Perhaps transforming or re-expressing the data in a more appropriate metric
(8) Selecting a computational procedure
(9) Carrying out the calculations
(10) Checking the calculations for reasonableness
(11) Interpreting the results

It is my view that the goal of a basic statistics course should be to help students understand as much as possible of the statistical analysis process. They will need to understand this process before they can begin to use statistics, of course. But, just as importantly, they will need to understand it if they are to be intelligent consumers of statistical results. Only by understanding some of what should be done, will they have any hope of being able to judge whether a given analysis has been done appropriately or not.

The concepts involved in this analysis process are apparently not very easy to most students. This is especially true if they become ensnarled in a mass of computational details before they begin to get an understanding of the process as a whole.

Computers and calculators can help some, but they do not provide a complete solution by any means. Note in particular that pocket calculators can only help in steps (7) and (9). In fact, even computers can only help in steps (3), (6), (7) and (9). The rest must come from hard thinking. The best that any computational and plotting aid can do is to free the analyst's/student's mind from the need to worry about the details of the calculations or plots.

The important decisions of what to do are left to the analyst/student. But if computers have freed the teacher from the need to teach computational details, class attention can be focused instead on the more important concepts of the statistical analysis process.

POSSIBILITIES VERSUS TRADITION

In the past most attention in basic statistics courses has been focused on either calculations or on the properties of basic statistical formulas under the narrow assumptions of simple random sampling from a Gaussian distribution. It is now possible to change this emphasis and bring our teaching much more in line with how good statistical analysis is actually done.

Let us now consider the possibilities the computer affords us at the various steps in the analysis process.

(1) Understanding the context of the data. By using computers, more realistic data sets can be considered. They need not be made up, preselected or modified to make it easy for the calculations to be carried our nor to satisfy the narrow assumptions of random samples from Gaussian distributions. A richer variety of larger data sets in areas of interest to the students can be used.

(2) Understanding the goals of the analysis. Along with more realistic data sets comes a much broader class of possible goals. Thus the importance of choosing statistical procedures that are appropriate to the particular goals of the analysis can be emphasised.

(3) Checking the data. Almost every real data set has errors of the non-statistical sort. Dials are misread, impossible answers are recorded and so on. Learning how to look for these trouble spots with histograms, plots and frequency tables is made relatively easy by computers. It is also easy to dramatise the importance of such checking.

(4) Might data shed light on stated goals? More realistic problems can encourage students to think about how hard it really is to gather good data on many questions of interest. This is very important to their future roles as consumers of statistical analysis either in their everyday thinking or in their jobs.

(5) Sources of uncertainty. This again relates to the need to go beyond the assumption of a simple random sample. In many cases the data are, in no sense, any sort of random sample. We can, at best, use the results of standard procedures for guidance (see, e.g. Joiner, 1982; Tukey, 1982). In other cases more complex variance structures are present. Understanding these structures, and their impact on standard procedures is quite important.

(6) Making a variety of plots. Good plots are often very important to the statistical analysis process and in some cases even obviate the need for a more formal analysis. It is here that computers really excel. Plotting by hand a set of 100 observations is a very tedious, error prone activity. Few students or teachers would make more than a handful of such plots by hand in a course. Yet good practice often calls for a number of such plots in each analysis. Data must be plotted in a variety of ways; residuals from various analyses must be plotted in a variety of ways; subsets of data must be plotted; and so on. (See, for example, Joiner, 1981a).

(7) Transformations of data. Data are not necessarily recorded in the units most appropriate for their analysis and much can often be gained from a change of units: gasoline mileage may be better analysed in gallons per mile than in miles per gallon; acidity in pH rather than in hydrogen ion concentration; toxicity in deaths per hour rather than in hours to death; stock prices in earnings per cost of stock rather than price earnings ratio; and so on. Computers make it easy to re-express data then replot and recompute. (See, e.g. Joiner, 1981b). And, they make these processes seem quite easy and natural.

(8) Selecting a computational procedure. Computers are of little help here, except that by making it easy to do a variety of calculations, they may make it more likely that a procedure will be chosen based on its appropriateness rather than on its ease of calculation. One minor problem is that computers make it almost too easy to do more analyses than are justified. Students should be warned against the trap of doing analysis after analysis until they find one that gives "statistically significant" results.

(9) Carrying out the calculations. Computers greatly facilitate this process in two ways. They obviously make the numerical calculations quite easy, but in addition, their ability to store a copy of the carefully checked data frees the student/analyst from the need to re-enter the data. They also make it easy to loop back and begin again at step (6) by plotting the data in new and more effective ways, and possibly retransforming the data, for another analysis.

(10) Checking calculations for reasonabiness. After getting an answer it is important to reflect on whether that answer is reasonable or not given what else we know about the problem and the data. Computers are presently of little help here except in providing more time for this activity. One way they might help, although this seems to be not much exploited, is in providing drill exercises in which students are presented with a list of numbers or a histogram then asked to guess some characteristic of the data such as their standard deviation.

(11) Interpreting the results. Again computers are of little help except in freeing up time to concentrate on this important activity and in making it easy to do further numerical or graphical processing of results. For example, they make it easy to compute and plot predicted values and residuals.

SUMMARY

Computers can be a big help to us in teaching statistics. They can free us of the need to focus on computational details and narrow, unrealistic assumptions. To take full advantage of this opportunity we will need to free our minds of what we have here-to-fore taken for granted must be taught. We must do fresh thinking and study how good statistical design and analysis are actually done in practice - then see how we can best teach this process to our students. If we do this, I think our students will find statistics a much more interesting and relevant subject and will become better users of statistical procedures and better consumers of statistical analyses. We will also have more fun teaching it.

Acknowledgements

I am very grateful to Alison K. Pollack who made a number of suggestions which significantly improved this report.

References

[1] Joiner, Brian L. (1981a) "Lurking Variables: Some Examples". *The American Statistician*, 35, 227-233.

[2] Joiner, Brian L. (1981b) "How To Use Minitab Effectively in Teaching Statistics". *Proc. Gothenburg Symposium on Teaching Statistics*, Lennart Råde (ed.) 16-30.

[3] Joiner, Brian L. (1982) "Practicing Statistics or What They Forgot to Say in the Classroom", *Teaching of Statistics and Statistical Consulting*, Jagdish, S. Rustagi and Douglas A. Wolfe (eds.) Academic Press, New York, 327-342, 391-393.

[4] Koopmans, Lambert H. (1981) An Introduction to Contemporary Statistics, Duxbury Press, Boston, MA.

[5] Ryan, Thomas A. Jr., Brian L. Joiner and Barbara F. Ryan (1976) Minitab Student Handbook, Duxbury Press, Boston, MA.

[6] Ryan, Thomas A. Jr., Brian L. Joiner and Barbara F. Ryan (1982) Minitab Reference Manual, Minitab Project, University Park, PA.

[7] Tukey, John W. (1977) Exploratory Data Analysis, Addison-Wesley, Reading, MA.

[8] Tukey, John W. (1982) "Discussion of Papers by Calvin, Joiner and van Belle", Teaching of Statistics and Statistical Consulting, Jagdish S. Rustagi and Douglas A. Wolfe (eds.) Academic Press, New York, 379-389.

COMPUTER ALGORITHMS FOR THE CLASSICAL DISTRIBUTION FUNCTIONS

H O Posten, University of Connecticut

INTRODUCTION

Access to a collection of computer programs for the classical distribution functions, particularly if these can be readily used interactively by any student, is an important teaching tool for teachers of probability and statistics. The availability of such a collection allows the teacher and student to concentrate on the concepts and development of probability and statistics with little if any emphasis on the details of determining probabilities from tables.

By using the algorithms to be presented, any teacher with beginning level programming ability (or with the assistance of a colleague or student with this ability) can provide a full set of programs for evaluating binomial, Poisson, geometric, negative binomial, hypergeometric, normal, *t*, *F*, beta, gamma, and chi-square probabilities. In most cases, the actual algorithm will be reasonably simple and, for the continuous distributions, general enough to allow a user to increase accuracy by specifying the number of terms or steps to be used. Also, more than one algorithm for each distribution will be provided to allow a user to check the quality of one algorithm with another.

ALGORITHMS FOR DISCRETE DISTRIBUTIONS

Recursive Algorithm

A single recursive algorithm is very effective for evaluating probabilities for any of the discrete distributions previously mentioned. This algorithm is based on a common property of each of these distributions, that

$$f(x) = c(x-1)f(x-1) \tag{1}$$

for all possible values of x appropriate to each distribution, where $f(x) = P(X=x)$ is the probability function and $c(x-1)$ is a simple multiplicative factor. The distribution function $F(x) = P(X \leq x)$ is evaluated by accumulating values of $f(x)$. For each distribution, the appropriate details are provided in Table 1. The geometric distribution is not provided in Table 1 since it is a special case of the negative binomial distribution with $r = 1$.

From Table 1 and formula (1), the general recursion algorithm is given in Table 2. This algorithm calculates the first probability $f(0)$ and, from this, each successive probability $f(k)$ until the probability for the desired value of x is reached. The individual probabilities $f(k)$ are accumulated at each step to provide the cumulative probability. This algorithm, as listed in Table 2, provides only the values of f(x) and F(x) but can easily be modified to efficiently produce a complete table of probabilities. To do this, it is necessary to output the values of f(k) and F(k) for each value of k in steps 2 and 3 and to continue step 3 until the largest value of k is reached for the given distribution. For the two infinite range distributions, there is no largest value for k and the probability calculations must be arbitrarily terminated at some point to avoid continuing the calculations forever. A convenient strategy, in this case, is to stop the calculations when essentially all of the probability has been calculated. That is, prespecify an allowable RESIDUAL amount of probability and terminate the calculations when the accumulated probability exceeds 1-RESIDUAL. By specifying RESIDUAL to be quite small, all but a negligible amount of probability will be provided in the resulting table. Table 3 provides the modified algorithm for producing a table of probabilities for any of the listed discrete distributions.

TABLE 1: Elements of Recursive Algorithm for Discrete Distributions

distn.	f(x)	range of x	c(x-1)	f(0)
binomial	$\binom{n}{x}p^x q^{n-x}$	x = 0,1,...,n	$\frac{p(n-x+1)}{qx}$	q^n
Poisson	$\frac{e^{-m}m^x}{x!}$	x = 0,1,2,...	$\frac{m}{x}$	e^{-m}
negative binomial	$\frac{(r+x-1)!}{(r-1)!x!}p^r q^x$	x = 0,1,2,...	$\frac{(r+x-1)q}{x}$	p^r
hyper-geometric	$\frac{\binom{M1}{x}\binom{M2}{n-x}}{\binom{M1+M2}{n}}$	usually x = 0,1,...,n	$\frac{(M1-x+1)(n-x+1)}{x(M2-n+x)}$	$\frac{\binom{M2}{n}}{\binom{M1+M2}{n}}$

TABLE 2: Recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Probabilities

STEP 1: Input value of x and the distribution parameters (e.g., n and p for the binomial).

STEP 2: Evaluate f(0) and F(0) = f(0).

STEP 3: Evaluate f(k) = c(k-1)f(k-1) and F(k) = F(k-1) + f(k) for k = 1,2,...,x.

STEP 4: Output x, f(x), F(x).

TABLE 3: Recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Tables

STEP 1: Input distribution parameters (and RESIDUAL for an infinite range distribution).

STEP 2: Evaluate f(0) and F(0) and output 0, f(0), F(0).

STEP 3: Evaluate $f(h) = c(h-1)f(h-1)$ and $F(h) = F(h-1) + f(h)$ and output h, f(h), F(h) for $h = 1,2,3,\ldots,d$, where d is the largest possible value for h.

STEP 3* (replaces STEP 3 when distribution range is infinite): Evaluate $f(h) = c(h-1)f(h-1)$ and $F(h) = F(h-1) + f(h)$ and output h, f(h), F(h) for $h = 1,2,\ldots$ until $F(h) > 1-\text{RESIDUAL}$.

Except for hypergeometric distributions, the calculation of f(0) for the algorithms in Tables 2 and 3 is usually quite simple using the forms given in Table 1. The calculation of f(0) for the hypergeometric distribution, however, presents some difficulties since f(0) is the ratio of two combinatorials each involving factorials. A special approach to the evaluation of f(0) is therefore needed for the hypergeometric distribution. It is easy to show for this distribution that f(0) may be written as

$$f(0) = \prod_{j=1}^{n} \frac{M2-j+1}{M1+M2-j+1} , \qquad (2)$$

and a simple program loop can be used to evaluate f(0) as a *product of ratios*.

The foregoing indicates that it is not difficult to write a recursive program to evaluate any of the indicated distributions with the use of (1), even for the hypergeometric distribution. For the latter distribution, this holds provided the range is $x = 0,1,2,\ldots,n$. Occasionally, one runs into a hypergeometric distribution for which the sample size is larger than one or both of M1 and M2. In this case, at least one end of the range is restricted. If a general hypergeometric program is required to handle this case, as well as the usual case, the programming is a bit more difficult. It must be taken into account that the general range for the hypergeometric random variable is $x = c, c+1, \ldots, d$, where $c = \max(0, n-M2)$ and $d = \min(n, M1)$. The lowest possible value of x may, therefore, be zero or larger than zero. In the latter case, the algorithms of Tables 2 and 3 must be modified so that f(0) is replaced in STEP 2 by f(c) when $c > 0$, and in STEP 3, h should start at $c+1$. It is not difficult to show that (when $n > M2$) f(c) may be written as

$$f(c) = \prod_{j=1}^{M2} \frac{n-j+1}{M1+M2-j+1} \qquad (3)$$

The form (3) makes it easy to evaluate f(c) as a product of ratios. Also, when producing a table for the hypergeometric distribution, the calculations in STEP 3 (Table 3) should be terminated when h reaches the maximum value min(n,M1). For this much generality, some additional programming effort is required. The result is worth the effort, however, since any hypergeometric problem can be handled without tables.

In general, the foregoing recursive method accurately evaluates probabilities for any of the listed discrete distributions, even when the range of x is large. It is fairly easy, using this algorithm, to achieve at least ten decimal accuracy by carrying enough digits in the calculations (for example, 16 decimals in double precision calculations on an IBM 360/370 System).

A Warning and a Modification

On reflection, the foregoing recursive algorithm is seen to have one flaw. Should f(0) have a value smaller than the smallest magnitude allowed by the computer, underflow will occur and f(0) will be read as zero by the computer. In this event, all subsequent probabilities f(k) will be calculated as zero since each is the product of a multiplier and the previously calculated probability. (Such an error can be avoided by incorporating into the program at STEP 2 of the algorithm a statement that stops the program when f(0) is read as a zero and informs the user to use another method). In this case, for the finite range random variables one can reverse the recursive algorithm starting at the highest possible value of the random variable and accumulate probabilities backwards instead of forwards. It is possible that both the first and last probabilities can be read as zero because both are too small, but many times this backwards approach can be useful in sidestepping the problem of f(0) being too small.

In practice, the problem of f(0) being too small does not seem to arise for the range of hypergeometric parameter values usually met. Also, for the negative binomial distribution, it is unlikely that this problem would occur for the range of p and r usually met by teachers and students. For this to occur, p would have to be particularly small and/or r would have to be unusually large. The problem does occur for the binomial and Poisson distributions (for the Poisson, $m > 180$ and, for the binomial with $p = .50$, $n > 260$ will yield an f(0) that is too small for an IBM 360/370 System). Consequently, it is wise to provide a program termination when such a problem occurs and then use a backup program. An effective backup procedure for such cases is a modification of the foregoing recursive algorithm which uses a natural logarithm and an exponential operation at each step. Specifically, from (1),

$$\ln f(x) = \ln c(x-1) + \ln f(x-1) \qquad (4)$$

for all appropriate values of x for each of the listed distributions. The modification evaluates $\ln f(0)$ ($\ln f(0) = n \ln q$ for the binomial and $\ln f(0) = -m$

for the Poisson distribution) and subsequently $\ell n\ f(k)$ from this using (4) repeatedly. The values of $f(0)$ and $f(k)$ are obtained by exponentiating these natural log values. This avoids the problem of $f(0)$ being too small for the computer for any of the four discrete distributions in Table 1. The appropriate algorithms are provided in Tables 4 and 5.

From a teaching viewpoint, the author prefers to use the recursive algorithm in its simplest form with a terminating statement that informs the user to use another method when $f(0)$ is read as zero. In this form, the algorithm is very simple and can readily be explained to anyone with algebra background, including high school students. As a backup method, for the binomial and Poisson, programs using the log-recursive algorithm are quite effective, ordinarily achieving at least 10 decimal accuracy when enough digits are carried throughout the calculations. Such a backup program is unnecessary for the hypergeometric distribution and probably for the negative binomial, though either can be developed from Tables 4 and 5.

TABLE 4: Log-recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Probabilities

---

STEP 1: Input value of x and the distribution parameters.

STEP 2: Evaluate $\ell n\ f(0)$ (= $n\ \ell n\ q$ for binomial and $-m$ for Poisson), $f(0) = \exp(\ell n\ f(0))$, and $F(0) = f(0)$.

STEP 3: Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, and $F(k) = F(k-1) + f(k)$ for $k = 1,2,\ldots,x$.

STEP 4: Output x, $f(x)$, $F(x)$.

TABLE 5: Log-recursive Algorithm for Binomial, Poisson, Negative Binomial, and Hypergeometric Tables

---

STEP 1: Input distribution parameter values (and RESIDUAL for an infinite range distribution).

STEP 2: Evaluate $\ell n\ f(0)$, $f(0) = \exp(\ell n\ f(0))$, $F(0) = f(0)$ and output 0, $f(0)$, $F(0)$.

STEP 3: Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, and $F(k) = F(k-1) + f(k)$ and output k, $f(k)$, $F(k)$ for $k = 1,2,\ldots,d$, where d is the largest possible value for k.

STEP 3* (replaces STEP 3 when distribution has an infinite range): Evaluate $\ell n\ f(k) = \ell n\ c(k-1) + \ell n\ f(k-1)$, $f(k) = \exp(\ell n\ f(k))$, $F(k) = F(k-1) + f(k)$ and output k, $f(k)$, $F(k)$ for $k = 1,2,\ldots$ until $F(k) > 1\text{-RESIDUAL}$.

Direct Term by Term Evaluation of Individual Probabilities

A relatively simple algorithm for any of the distributions in Table 1 evaluates the probability $f(x)$ by evaluating each factor and multiplying these together. Such individual probabilities can then be added for cumulative probabilities.

This approach is not usually practical because the combinatorials involved in each of the f(x) can become too large for the computer. This approach, however, becomes practical if logarithms are used and one has an accurate algorithm for evaluating the logarithm of a factorial. The natural logarithm of a factorial can be evaluated using $\ln x! = \ln \Gamma(x+1)$, where $\Gamma(y)$ is the gamma function. An excellent algorithm for evaluating $\ln \Gamma(y)$ will be provided later on in the section headed 'Algorithms for Evaluating the t- and F-Distribution Functions'. The general algorithm is simple and will be illustrated using the binomial distribution. For this distribution,

$$\ln f(k) = \ln \Gamma(n+1) - \ln \Gamma(k+1) - \ln \Gamma(n-k+1) + k \ln p + (n-k) \ln q. \tag{5}$$

The value of $f(k) = \exp(\ln f(k))$ is obtained from (5) for all required values of $k = 0,1,\ldots,x$ and accumulated to provide the cumulative probability $F(x)$. Corresponding algorithms for the other three discrete distributions are obtained similarly. This method is slower than the two recursive methods, but it is effective and can achieve 10 or more decimal place accuracy (carrying 16 decimals in the computer). For the evaluation of a single probability, $f(x)$, it is relatively fast and most serviceable. For cumulative probabilities it is useful as an independent algorithm for checking the quality of the two recursion algorithms. For teaching purposes, the latter are better since they are faster and do not involve factorials, thereby allowing easier explanation.

Other Methods for Discrete Distributions

A fourth effective method for evaluating a cumulative binomial probability is bsed on the identity

$$F(x;p,n) = \sum_{k=0}^{x} \binom{n}{x} p^x q^{n-x} = 1 - I_p(x+1,n-x) , \tag{6}$$

where

$$I_y(a,b) = \Gamma(a+b) \, [\Gamma(a)\Gamma(b)]^{-1} \int_0^y t^{a-1}(1-t)^{b-1}dt \tag{7}$$

is the cumulative beta distribution function. The identity (6) is given by various authors including Greenwood and Hartley (1962) and Kennedy and Gentle (1980). With an accurate algorithm for evaluating the beta distribution function, (6) provides a good algorithm for evaluating the cumulative binomial distributive. Three good algorithms will be provided in the section 'Algorithms for the Beta Distribution Function' for evaluating the beta distribution function. Using these beta algorithms, very good accuracy (ten or more decimals, if enough digits are carried in the calculations) can often be achieved. In fact, using the three methods for evaluating the cumulative binomial probability (recursive, log-recursive, and binomial-beta) often leads to 10 or more decimal place agreement between the three methods (using double precision on an IBM

360/370 computer).

A third effective method for evaluating a Poisson cumulative probability is based on the identity

$$F(x;m) = \sum_{k=0}^{x} \frac{e^{-m}m^{k}}{k!} = 1 - G_{2m}(x+1,2) \tag{8}$$

given by Greenwood and Hartley (1962) and by Kernedy and Gentle (1980) where

$$G_{y}(a,b) = [\Gamma(a)b^{a}]^{-1}\int_{0}^{y} t^{a-1}e^{-t/b}\,dt \tag{9}$$

is the gamma distribution function. The quality of this Poisson algorithm will depend upon the quality of the gamma algorithm. This, however, will present no problem since three excellent gamma algorithms will be provided in the section 'Algorithms for the Gamma Distribution Function'.

ALGORITHMS FOR CONTINUOUS DISTRIBUTIONS

Algorithms for the Normal Distribution

Probably the most effective method for evaluating a single standard normal probability is a continued fraction expansion. Such an approach is given by Bargmann (1969) as

$$\Phi(z) = 1 - \phi(z)CF \tag{10}$$

where $\Phi(z)$ is the standard normal distribution function, $\phi(z)$ is the corresponding density function,

$$CF = \frac{z}{1-}\,\frac{z^2}{3+}\,\frac{2z^2}{5-}\,\frac{3z^2}{7+}\,\frac{4z^2}{9-}\,\cdots \quad \text{for } 0 < z \leq 2.5 \tag{11}$$

and

$$CF = \frac{1}{z+}\,\frac{1}{z+}\,\frac{2}{z+}\,\frac{3}{z+}\,\frac{4}{z+}\,\cdots \quad \text{for } z > 2.5 \tag{12}$$

The expansions (11) and (12) are continued fraction expansions which are truncated (i.e., terminated at some step) for evaluation and are then rational functions. For example, (11) is

$$CF = \cfrac{z}{1-\cfrac{z^2}{3+\cfrac{2z^2}{5-\cfrac{3z^2}{7+4z^2}}}}$$

if the truncation is made at the indicated step. It is often convenient to evaluate continued fractions recursively. In this direction, (11) may be evaluated by the steps in Table 6 and (12) by the steps in Table 7. Tables 6 and 7 provide a quick and easily programmed method for evaluating CF. The cumulative normal probability, $\Phi(z)$, is then readily calculated by (10).

TABLE 6: Recursive Evaluation of Normal Continued Fraction ($0 < z \leq 2.5$)

STEP 1: Evaluate for prespecified N, $W(N) = 2N-1+N(-1)^N z^2$.

STEP 2: For K = N-1,N-2,...,1, evaluate $W(K)=2K-1+K(-1)^K z^2/W(K+1)$.

STEP 3: Evaluate $CF = z/W(1)$.

TABLE 7: Recursive Evaluation of Normal Continued Fraction ($z > 2.5$)

STEP 1: Evaluate for prespecified N, $W(N) = z + N/z$.

STEP 2: For K = N-1,N-2,...,1, evaluate $W(K) = z + K/W(K+1)$.

STEP 3: Evaluate $CF = 1/W(1)$.

A careful study by one of my students, P. Hormel, for z = 0.1 to 10.0 by .10 showed that less than 40 erms in the continued fraction were necessary to provide at least 13 decimal accuracy for all z in the range $0 < z < 10$, using double precision calculations on an IBM 360/370 computer.

A second algorithm which also provides very good accuracy is the series expansion (13) provided by Abramowitz and Stegun (1964).

$$\Phi(z) = \frac{1}{2} + \phi(z) \sum_{k=0}^{\infty} T_k, \tag{13}$$

and

$$T_k = \frac{z^{2k+1}}{1(3)(5)\ldots(2k+1)}. \tag{14}$$

The normal distribution function, $\Phi(z)$, is evaluated with a truncation of the sum in (13), which itself can be evaluated using the recursive relation

$$T_0 = z, \; T_k = \frac{z^2}{2k+1} T_{k-1}, \; k = 1,2,\ldots . \tag{15}$$

The specific algorithm is given in Table 8. because of teh recursive element, this algorithm is reasonably fast. It is also quite accurate. The study by P. Hormel showed that 100 terms will provide at least 12 decimal accuracy for $0 < z < 9.3$ and at least nine decimal accuracy for values of z up to z = 10.

A third very effective method for evaluating the normal distribution function is the simple method of numerical integration known as Simpson's method. This method divides the range of integration into N intervals of length 2H and then approximates the area over each interval by a simple quadratic approximation to the normal density function. The specific algorithm for evaluating a cumulative standard normal probability, from $z = L$ to $z = U$, is

$$P(L < z < U) = \sum_{k=1}^{N} A_k \quad . \tag{16}$$

where

$$A_k = (H/3)[\phi(z_{2k}) + 4\phi(z_{2k-1}) + \phi(z_{2k-2})]$$

is the approximate area over the range z_{2k-2} to z_{2k} and $z_m = L+mH$. The approximations A_k are then accumulated. This method is surprisingly effective for the normal distribution, even with relatively large values of H, but the most effective value of H (using double precision on an IBM 360/370) is about $H = .005$. With this value of H, accuracy of at least 10 decimal places is obtained for all values $0 < z < 10$.

TABLE 8: Series Algorithm for the Normal Distribution Function

```
STEP 1:  Input $z$ and $N$ = number of terms to be used in the series
         in (13).
STEP 2:  Evaluate $T_0 \equiv z$, $SUM = T_0$.
STEP 3:  For $k = 1,2,\ldots, N$, evaluate $T_k = [z^2/(2k+1)]T_{k-1}$ and
         $SUM = SUM + T_k$.
STEP 4:  Evaluate $\Phi(z) = (1/2) + \phi(z)\, SUM$ and output $z$, $\Phi(z)$.
```

Of the three methods, the most effective for evaluating a single normal probability is the continued fraction algorithm because of its high accuracy when using only a small number of terms in the continued fraction. However, the Simpson method algorithm is the most effective for producing a table of standard normal probabilities, sicne the area, A_k, over each of the N intervals is directly calculated and accumulated. The evaluation of each A_k is fast, and it is a simple matter to output accumulated areas as each A_k is calculated to provide a table. Students find it interesting to produce their own normal tables which are better than those in any book. The programming challenge is to print out the entire table on one or two regular computer pages (11 by 14 inches). A table of cumulative standard normal probabilities for $z = 0.00$ to 7.00, by 0.01, can be organized in 70 lines with 10 probabilities in each row, each provided to nine decimal places.

Algorithms for Evaluating the t- and F-Distribution Functions

The distribution functions for the t- and F-variates are quite simple to evaluate, provided an algorithm is available for the beta distribution function $I_y(a,b)$ defined by (7). For the t-variate with v degrees of freedom, the identity

$$P(t_v < t) = (1/2)[1 + I_x(1/2, v/2)], \text{ for } t > 0, \tag{17}$$

with

$$x = t^2/(v+t^2)$$

provides the algorithm. For $F(v_1,v_2)$, the F-variate with v_1 and v_2 degrees of freedom, the identity

$$P(F(v_1,v_2) < F) = I_x(v_1/2, v_2/2) \tag{18}$$

with

$$x = (v_1/v_2)F/[1 + (v_1/v_2)F]$$

yields the algorithm. The identities (17) and (18) are given by Johnson and Kotz (1970).

Algorithms for the Beta Distribution Function

An effective way to evaluate a beta distribution function is to use a continued fraction expansion. One such algorithm, due to Muller (1931) and given by Abramowitz and Stegun (1964), Bargmann and Ghosh (1963), Boardman and Kopitzke (1975) and Kennedy and Gentle (1980), is

$$I_x(a,b) = \frac{\Gamma(a+b)}{\Gamma(a+1)\Gamma(b)} x^a (1-x)^{b-1} CF \tag{19}$$

where

$$CF = \frac{1}{1-}\,\frac{c_1}{1+}\,\frac{d_1}{1-}\,\frac{c_2}{1+}\,\frac{d_2}{1-}\cdots \tag{20}$$

is a continued fraction with coefficients

$$c_k = \frac{(b-k)(a+k-1)x}{(a+2k-2)(a+2k-1)(1-x)}\ , \quad d_k = \frac{k(a+b+k-1)x}{(a+2k)(a+2k-1)(1-x)}\ . \tag{21}$$

One problem with this method is the requirement to evaluate the gamma functions in (19). This can be handled by using an efficient algorithm for evaluating the logarithm of a gamma function. An excellent procedure is provided in the section 'Algorithm for the Natural Logarithm of the Gamma Function'. For the evaluation of (20), it is effective to use a recursive form, as given in Table 9.

TABLE 9: Continued Fraction Algorithm for the Beta Distribution Function

STEP 1:	Input parameter values a and b and the values of x and n = the number of recursive steps to be used (the actual number of steps used in the continued fraction will then be 2n).
STEP 2:	Evaluate $S_n = c_n/(1+d_n)$, using (21) to evaluate c_n and d_n.
STEP 3:	For $k = n-1, n-2, \ldots, 1$, evaluate $S_k = c_k/[1+d_k/(1-S_{k+1})]$, using (21) to evaluate c_k and d_k.
STEP 4:	Evaluate $CF = 1/(1-S_1)$ and $LA = \ln\Gamma(a+b) - \ln\Gamma(a+1) - \ln\Gamma(b)$.
STEP 5:	Evaluate $LB = LA + a\ln x + (b-1)\ln(1-x) + \ln CF$ and $I_x(a,b) = \exp(LB)$.

Bargmann and Ghosh indicate that this method is effective over a wide range of values of a and b, providing accuracy of at least 10^{-8} using 80 terms in the continued fraction (n = 40 in Table 9). However, Kennedy and Gentle recommend the use of this method for $x < (a-1)/(a+b-2)$ and the use of

$$I_x(a,b) = 1 - I_{1-x}(b,a) \tag{22}$$

for larger x. The present author's experience with this algorithm indicates that often 10 or more decimal place accuracy is achieved using significantly less than 80 continued fraction terms, even above the range for x given by

Kennedy and Gentle. Also, a study by one of my students, J. Leitz, over 91 beta distributions with $a \geq b$ and $0.1 \leq a,b \leq 500$ and over 15 x-values in the probability range 0.0001 to 0.9999 for each distribution, supports this conclusion. Leitz's study, though, suggests that this continued fraction algorithm may not be very effective for the case when a and b are both less than unity, as also suggested by Bargmann and Ghosh. An exception to this is the case $a = b = 0.5$, when the method is quite accurate (10 or more decimals).

Another effective algorithm for evaluating the beta distribution function is based on a Fourier series Chebyshev polynomial expansion (Woods and Posten (1977)). Computationally, the algorithm is effective because a recursive strategy is used with the expansion

$$I_x(a,b) = 1 - (\theta/\pi) - S \tag{23}$$

where

$$\theta = \cos^{-1}(2x-1) \quad \text{and} \quad S = \sum_{j=1}^{\infty} b_j \sin(j\theta) \; , \tag{24}$$

where the b_j coefficients will be defined. The series S is evaluated by truncating

$$S_n = \sum_{j=1}^{n} b_j \sin(j\theta) \tag{25}$$

at some convenient value for n. The coefficients are evaluated recursively by

$$\begin{aligned} b_1 &= (2/\pi)(a-b)(a+b)^{-1} \\ b_2 &= (1/\pi)(a+b)^{-1}(a+b+1)^{-1}[2(a-b)^2-(a+b)(a+b-1)] \\ b_{j+2} &= (a+b+j+1)^{-1}[2(a-b)\tfrac{j+1}{j+2}\, b_{j+1} + (j+1-a-b)\tfrac{j}{j+2}\, b_j)], \\ & \qquad j = 1,2,\ldots,n-2. \end{aligned} \tag{26}$$

S_n is then evaluated recursively and rapidly (without use of any trigonometric functions) by

$$\begin{aligned} U_{n+1} &= U_{n+2} = 0 \\ U_k &= b_k + (4x-2)U_{k+1} - U_{k+2} \; , \; k = n,n-1,\ldots,1 \\ S_n &= 2U_1 \, (x-x^2)^{1/2} \; . \end{aligned} \tag{27}$$

The algorithm is given in Table 10.

Studies by several students (F. Rockhold, A. Tupek and T. Ashbee) and myself indicate that this algorithm is effective over a broad range of values of a, b and x, achieving eight or more places if enough terms are used. Usually 200 terms provide at least 6-8 decimal accuracy and 500 terms, at least 8-10 decimal accuracy. Because of the recursive approach these terms are computed rapidly. The method is not as accurate as indicated if the distri-

TABLE 10: Fourier Series Chebyshev Polynomial Algorithm for the Beta Distribution Function

STEP 1:	Input values of a, b, x, and n = number of terms to be used in the series (25)
STEP 2:	Evaluate b_k, $k = 1,2,\ldots,n$, recursively from (26).
STEP 3:	Evaluate U_1 recursively from (27), and S_n from (27).
STEP 4:	Evaluate $I_x(a,b) = 1 - (\theta/\pi) - S_n$.

bution shape is such that the probability is strongly compressed towards either endpoint (*e.g.*, $a = 100$, $b = 2$). In this case it appears that one can usually expect four or five decimal accuracy with 200 terms. Additionally, this method is quite weak when either parameter is very near zero (a or $b < 0.2$) and is not very strong when either parameter is small ($0.2 \leq a$ or $b \leq 1.4$), though in the latter case one can anticipate several decimal place accuracy. This last weakness is an important consideration since, as indicated in the section 'Algorithms for Evaluating the t- and F-Distribution Functions', F-probabilities are evaluated using a beta algorithm. The beta parameters are each one-half of the corresponding degrees of freedom of the F. Thus, for example, $a = 1/2$ and $a = 1$ are important parameter values since they correspond to F-variates with one and two degrees of freedom. Fortunately, the general weakness for small parameter values does not exist when a or $b = 1/2$ (unless the other parameter value is a different low value, ≤ 1.4). The opposite is, in fact, true. When a or $b = 1/2$, the method has very high accuracy. On the other hand, when a or $b = 1$, the method is weak often providing as little as three decimal accuracy. In this case (and other cases), one can use an identity which evaluates the probability after first incrementing the low parameter. The following forms, provided by Abramowitz and Stegun (1964), increment the first parameter by one each time they are used:

$$I_x(a,b) = [\Gamma(a+b)/\Gamma(a+1)\Gamma(b)]x^a(1-x)^{b-1} + I_x(a+1,b-1) \tag{28}$$

$$I_x(a,b) = (1/x)[I_x(a+1,b) - (1-x)I_x(a+1,b-1)] \; . \tag{29}$$

Similar identities exist for incrementing the second parameter but (22) permits (28) or (29) to be used for this purpose.

A second continued fraction expansion can be used as a third method for evaluating the beta distribution function. This expansion, from Aroian (1941, 1959), is also given by Abramowitz and Stegun (1964), Bouver and Bargmann (1975), Boardman and Kopitzke (1975) and by Kennedy and Gentle (1980).

$$I_x(a,b) = \frac{\Gamma(a+b)x^a(1-x)^b}{\Gamma(a+1)\Gamma(b)} \left[\frac{1}{1-}\,\frac{c_1}{1+}\,\frac{d_1}{1-}\,\frac{c_2}{1+}\,\frac{d_2}{1-}\,\cdots\right] \tag{30}$$

with

$$c_k = \frac{(a+k)(a+b+k)x}{(a+2k)(a+2k+1)} \quad \text{and} \quad d_k = \frac{k(b-k)x}{(a+2k-1)(a+2k)} \; . \tag{31}$$

The algorithm, in recursive form, is identical to that in Table 9 using (30) and (31) in place of (19), (20) and (21) and noting that the exponent on (1-x) is increased by one in (30) as compared to (19) (i.e., b replaces b-1 in STEP 5). Bouver and Bargmann indicate that this continued fraction expansion provides a superior algorithm for evaluating the beta distribution function over an extensive range of values for a and b. They refer to the range $10^{-4} < a,b < 70{,}000$ and specifically state that the method is efficient even for extremely J-shaped and U-shaped distributions. They recommend its use for $x < a/(a+b)$ (i.e., for $x <$ mean), and the use of the complement form (22) for larger x.

Algorithms for the Gamma Distribution Function

Several algorithms for evaluating the gamma distribution function are quite effective over a wide range of parameter values. These algorithms can also be used to evaluate the chi-square distribution function, since the latter is a special case of the gamma distribution function

$$F(x;a,b) = [1/\Gamma(a)b^a]\int_0^x t^{a-1}e^{-t/b}\,dt \; , \tag{32}$$

with b = 2 and the degrees of freedom for the chi-square = 2a. The evaluation of (32) can be reduced to evaluating a gamma distribution function with b = 1,

$$I_x(a) = [1/\Gamma(a)]\int_0^x t^{a-1}e^{-t}\,dt, \tag{33}$$

since $F(x;a,b) = I_{x/b}(a)$. The algorithms to be presented will be for the form (33).

A very good continued fraction expansion for the gamma distribution, given by Khovanskii (1963), is

$$I_x(a) = [x^a e^{-x}/\Gamma(a)]\,\mathrm{CFK} \; , \tag{34}$$

where

$$\mathrm{CFK} = \frac{1}{a-}\,\frac{ax}{1+a+x-}\,\frac{(1+a)x}{2+a+x-}\cdots\frac{(n-1+a)x}{n+a+x-}\cdots \tag{35}$$

is a continued fraction expansion which can be evaluated recursively using CFK = 1/(a-T(1)), where

$$T(n) = \frac{(n+a-1)x}{n+a+x} \; , \; T(k) = \frac{(k+a-1)x}{k+a+x-T(k+1)} \; , \; k = n-1, n-2, \ldots, 1 \; . \tag{36}$$

The recursion algorithm is given in Table 11. Two of my students, K. Sullivan and M. El-Bolkiny, provided studies of this algorithm for 41 values of a in the range 0.05 to 200.25 over 18 x-values ranging in probability from 0.0001 to 0.999 or higher (except for a < 1, the lowest actual x-value was about x = 0.001). They found the Khovanskii method to be very effective, achieving at least 12 decimal place accuracy with at most 100 continued fraction terms (except for a > 100, up to 200 terms were required for this accuracy for some x-values).

A second continued fraction algorithm for the gamma distribution function is given by Boardman and Kopitzke (1975). They actually provide two continued fractions, one for lower x-values and the other for upper x-values. The two forms are

$$I_x(a) = [1/\Gamma(a+1)]x^a e^{-x}/CFL \; , \qquad (37)$$

where

$$CFL = 1 - \frac{x}{a+1+} \frac{x}{a+2-} \frac{(a+1)x}{a+3+} \frac{2x}{a+4-} \cdots \frac{(a+n-1)x}{a+2n-1+} \frac{nx}{a+2n-} \cdots \qquad (38)$$

and

$$I_x(a) = 1 - x^a e^{-x}[\Gamma(a)]^{-1}/CFU \; , \qquad (39)$$

where

$$CFU = 1 + \frac{1-a}{x+} \frac{1}{1+} \frac{2-a}{x+} \frac{2}{1+} \cdots \frac{n-a}{x+} \frac{n}{1+} \cdots \; . \qquad (40)$$

CFL and CFU are continued fraction expansions to be used for lower and upper values of x, respectively. Boardman and Kopitzke suggest using (37) when x < a and (39) otherwise. A recursive algorithm for evaluating CFL is

$$T(n) = x(1-n-a)/[a+2n-1+nx/(a+2n)] \qquad (41)$$

$$T(k) = x(1-k-a)/[a+2k-1+kx/(a+2k+T(k+1))], \quad k = n-1,\ldots,2,$$

$$CFL = 1-x/[a+1+x/(a+2+T(2))],$$

and for CFU,

$$T(n) = (n-a)/(x+n) \qquad (42)$$

$$T(k) = (k-a)/[x+k/(1+T(k+1))], \quad k = n-1,\ldots,2,$$

$$CFU = 1 + (1-a)/[x+1/(1+T(2))].$$

The recursive algorithm is provided by (37) and (41) for the lower expansion, and by (39) and (42) for the upper expansion. The pattern of the algorithm for each expansion is similar to that of the algorithm of Table 11.

Two of my students, E. Cook and G. Karbowski, provided studies of this algorithm over a set of 27 a-values in the range 0.1 ≤ a ≤ 200.25 and for seven values of x in the probability range 0.001 to 0.999 or higher (except for a < 1,

TABLE 11: Continued Fraction Algorithm for the Gamma Distribution Function

STEP 1: Input values of a, x and n = number of terms to be used in the continued fraction expansion.

STEP 2: Evaluation T(n) and T(k), k = n-1,...,1, from (36). Evaluate CFK = 1/(a-T(1)).

STEP 3: Evaluate $L = a \ln x - x - \ln \Gamma(a) + \ln CFK$ and $I_x(a) = \exp(L)$.

the lowest actual x was about x = 0.001). They found that the lower continued fraction was superior to the upper for the range suggested by Boardman and Kopitzke (and slightly beyond), except for a < 1 where CFL was found to be superior for x-values up to a probability level of about 0.999. Their recommendation was to use CFL for x-values up to a probability of 0.999 when a < 1 and up to a probability of 0.7 for $1 \leq a \leq 200.25$, and to use CFU elsewhere in this a-range. They found that this strategy provided accuracy of at least 12 decimals using at most 200 terms in the continued fraction. An exception to this was the largest a-value (a = 200.25) which required, for one x-value, 400 terms to achieve 12 decimal accuracy.

A third extremely effective algorithm for the gamma distribution function is the series method of Khamis (1965), also provided by Battacharjee (1970),

$$I_x(a) = \frac{x^a e^{-x}}{\Gamma(a)} \sum_{k=0}^{\infty} \frac{x^k}{a(a+1)\dots(a+k)} \quad , \tag{43}$$

for which the terms of the series can be evaluated recursively. If T_k represents the k-th term in the sum, $T_0 = 1/a$ and $T_k = (x/(a+k))T_{k-1}$ for k = 1,2,... . A study by M. El-Bolkiny over 41 a-values from 0.05 to 200.25 and 18 x-values in the range 0.0001 to 0.999 (except the lowest x was x = 0.001 for a < 1) showed this series algorithm to be surprisingly strong for all a and x in the range studied. At least 12 decimal accuracy was achieved in all cases using only 10 terms in the series. As a consequence, this algorithm seems superior to the previous two algorithms because of its efficiency and accuracy, at least over the range indicated. The appropriate algorithm is given in Table 12.

TABLE 12: Series Algorithm for the Gamma Distribution Function

STEP 1: Input values of a, x and n = number of terms to be used in the series expansion.

STEP 2: Evaluate $T_0 = 1/a$ and $SUM = T_0$.

STEP 3: For k = 1,2,...,n, evaluate $T_k = [x/(a+k)]T_{k-1}$ and $SUM = SUM + T_k$.

STEP 4: Evaluate $L = a \ln x - x - \ln \Gamma(a) + \ln SUM$ and $I_x(a) = \exp(L)$.

Algorithm for the Natural Logarithm of the Gamma Function

The best algorithm for evaluating $\ell n\,\Gamma(x)$ is probably the Euler-MacLauren expansion provided by Abramowitz and Stegun (1964).

$$\ell n\,\Gamma(x) = (x-1/2)\ell n\,x - x + (1/2)\ell n(2\pi) + \sum_{i=1}^{4} T_i \qquad (44)$$

with $T_1 = 1/12x$, $T_2 = -1/360x^3$, $T_3 = 1/1260x^5$, $T_4 = -1/1680x^7$.

This Euler-MacLauren expansion, truncated at four terms, provides an extremely accurate method for evaluating $\ell n\,\Gamma(x)$ with the error being less than 10^{-14} if $x \geq 18$ and enough digits are carried in the calculations. When $x < 18$, it is customary to increment the argument of the gamma function and use

$$\ell n\,\Gamma(x) = \ell n\;\Gamma(x+k) - \sum_{i=0}^{k-1} \ell n(x+i) \qquad (45)$$

where k is the integer which, when added to x, produces a value between 18 and 19. A program is easily written using (44) and (45).

CONCLUDING REMARK

The algorithms presented in this paper provide a useful array of methods for evaluating the classical distribution functions. The accuracy comments included were based on comments from other authors and, to a large degree, on personal experience. In the latter case, they should be understood to apply to the use of an IBM 360/370 computer system using double precision calculations.

References

[1] Abramowitz, M. amd Stegun, I.A. (1964) Handbook of Mathematical Functions, National Bureau of Standards, Applied Math. Series. Washington, D.C.: U.S. Govt. Printing Office.

[2] Aroian, L.A. (1941) Continued Fractions for the Incomplete Beta Function. Ann.Math.Statist., 12, 212-223; Correction (1959) 30, 1265.

[3] Bargmann, R.E. and Ghosh, S.P. (1963) Statistical Distribution Programs for a Computer Language. IBM Watson Research Center Report RC-1094.

[4] Bargmann, R.E. (1969) Principles and Design of Statistical Computer Languages, Chapter 8 in Progress in Operations Research, Vol.3, New York: John Wiley and Sons, Inc.

[5] Bhattacharjee, G.P. (1970) Algorithm AS 32: The Incomplete Gamma Integral. Applied Statistics, 19, 285-287.

[6] Boardman, T.J., and Kopitzke, R.W. (1975) Probability and Table Values for Statistical Distributions. Proc.Statist.Comput.Section. Washington, D.C.: American Statistical Association.

[7] Bouver, H. and Bargmann, R. (1975) Computer Algorithm and Modules for the Evaluation of Statistical Distribution Functions. Tech. Report 110, Dept. of Statistics, University of Georgia, Athens.

[8] Greenwood, J.A. and Hartley, H.O. (1962) Guide to Tables in Mathematical Statistics. Princeton, N.J.: Princeton University Press.

[9] Johnson, N. and Kotz, S. (1970) Continuous Univariate Distributions, Vol.2, Boston: Houghton Mifflin Co.

[10] Khamis, S.H. (1965) Tables of the Incomplete Gamma Function Ratio. Darmstadt, Germany: Justus von Liebig verlag.

[11] Kennedy, W.J. and Gentle, J.E. (1980) Statistical Computing. New York: Marcel Dekker, Inc.

[12] Khovanskii, A.N. (1963) The Application of Continued Fractions and their Generalizations in Approximation Theory. Gronigen, Netherlands: P. Noordhoff.

[13] Muller, J.H. (1931) On the Application of Continued Fractions to the Evaluation of Certain Integrals, with special Reference to the Incomplete Beta Function. Biometrika, 22, 284-297.

[14] Woods, J.D. and Posten, H.O. (1977) The Use of Fourier Series in the Evaluation of Probability Distribution Functions. Commun.Statist. B6, (3), 201-219.

USE OF HANDHELD PROGRAMMABLE CALCULATORS IN STATISTICS COURSES FOR NONTECHNICAL STUDENTS

D R Barr, Office of Naval Research, London Branch
(on leave from the Naval Postgraduate School, Monterey, California.

INTRODUCTION

The usefulness of stochastic models and statistical inferences in a wide range of disciplines has gradually led to the inclusion of probability and statistics courses in many academic curricula. Unfortunately, the required course in statistics is often the most disliked and feared course in a student's curriculum. This is especially the case for "nontechnical" students who generally have limited mathematical capabilities. Instructors of statistics courses for such students often feel they must devote a significant proportion of their instruction to improving the students' algebraic and computational skills. Clearly, allocating course time to algebra and arithmetic cuts into the time that can be spent on statistics.

It is reasonable, therefore, to examine the goals of the statistics courses for such nearly mathematically illiterate students. For many curricula, such examination reveals that the students need to gain an appreciation of statistical reasoning and applications, but it is not absolutely necessary that they be able to carry out the details of the associated computations. In their professional careers, these people will generally need to be informed consumers of statistics; for the most part they will never be skilled producers of statistics. But in order to learn about statistics, isn't it necessary for the student to "do" statistics? And if so, how can such a statistics course avoid the drudgery (from the student's point of view) of the algebra and calculation?

One possibility for teaching an applied statistics course without spending time on mathematical prerequisites is to use a computational device which allows the use of "canned" programs for carrying out the calculations. The particular device used for this purpose can take many forms, ranging from a programmable calculator to a microcomputer or a mini or mainframe computer. For planners of courses relying on such devices, there is an unfortunate tendency to get bogged down in considerations of the specific device to be used. The general approaches that can be used in teaching these courses is rather invariant with respect to the particular computational device used, as long as it is possible to use stored programs of sufficient sophistication. I shall return to this point at the conclusion of this paper.

In what follows, I shall report on a program we have been using for Management Science students at the Naval Postgraduate School in Monterey, California.

In this program, Texas Instruments TI59 calculators are issued to all students entering the Management Science (MS) curriculum, and these personal calculators are used by the students throughout their six quarter program leading to a Master's degree in Management Science. These students are military officers who received Bachelor's degrees in a wide variety of subject areas several years before entering the Master's degree program at the Naval Postgraduate School. These students are given a calculus "refresher" during their first quarter (a few of the students apparently haven't had a calculus course before). This is a calculator based course, using pre-recorded programs for computation of limits, derivatives, integrals and so on. An introduction to use of the TI59 calculator occupies the first week or so of this course. The two quarter sequence in probability and statistics follows the calculus refresher course; the role of the programmable calculator in this sequence is described below.

THE COURSE

The two quarter course in probability and statistics for the Management Science curriculum is in most respects a "standard" course for such students. The textbook currently used is Wannacott and Wannacott, "Introduction to Statistics" (Wiley). A written supplement was prepared which served as an interface between the text and the calculator. This supplement contained worked examples of problems as well as listings of the programs distributed to the students, and instructions for using the programs. The programs themselves were distributed in the form of pre-recorded magnetic strips. As each topic was encountered in the text, the corresponding supplement material and appropriate programs were discussed in the class and use of the programs was demonstrated by the instructor. Emphasis was placed on the program input requirements (statistical assumptions, data) and interpretation of the program outputs.

This approach allowed adherence to the standard course outline, while emphasizing the idea of making the students into informed consumers of statistics. Several rather unique consequences were observed in this use of computation devices as an integral course resource:

(i) Students are nearly always able to calculate the correct numerical answer to standard problems. When incorrect numerical answers did occur, they were due to erroneous data entry or, more seriously, inappropriate choice of program by the student. This tends to make a student's performance on a given problem type a Bernoulli variable, without much chance for assignment of partial credit. This led us to conduct the grading for the course along the lines of the "mastery concept". Indeed, a list of specific objectives were given to the students for each chapter of the course. In this way, the students knew exactly what they were expected to be able to do. Any student who "mastered" the approximately

200 objectives would receive an A in the course.

(ii) Students usually found calculations of solutions to problems in the text to be quite simple, and they often requested supplementary exercises. This placed something of a burden on the instructors.

(iii) There were occasional mechanical failures of the devices. I recall once, during an examination, there was a very brief power failure that caused the calculator memories to be cleared. There were occasional problems with individual calculators in getting a program card to read, having insufficient battery charge, and so on. Such failures are disturbing to the student because they can affect his (her) grade and yet are not the student's fault. I learned to bring one or two spare calculators into exams, which I could lend to students in case of such failures.

(iv) Occasionally, students just didn't like using the calculator. Such students attempted to solve the homework and examination problems by traditional hand calculation. While this is in some ways a good thing, it takes time; students insisting on hand calculation tended to receive lower course grades because they didn't have time to finish exams.

(v) For any course which includes discussion of programs and demonstration of their use in solving statistical problems, it is necessary that the students all have the same model of calculator. As mentioned above, we were able to arrange to have T159's distributed to all students in our course. This was done with a pool of calculators owned and maintained by the school. Students checked calculators out of this pool. An alternative would be to require that students purchased a specified calculator as a precondition for taking the course. This idea is becoming more feasible as the cost of sophisticated calculators falls. I recall that I was required to purchase a rather fancy slide rule when I first entered an engineering curriculum in the late 1950's. That slide rule cost a little over $30, which was roughly the cost of a modest bicycle at the time. The T159, for example, is available at about $170, which is roughly the cost of a modest bicycle today.

In addition to providing easy calculation of numerical results for exercises in the "traditional" text, the use of calculators allowed a number of non-traditional statistical computations to be performed by the students. Examples are: (a) simulations (both as a way of examining such ideas as confidence intervals, and as an analysis tool in itself), (b) robustness (for example, seeing how the posterior probabilities are affected by variations in the prior probabilities in an application of Bayes' formula), (c) significance testing (wherein the level of significance of the observed test statistic is calculated), and (d) applications of transformations (although transformations such as coding to simplify calculations aren't necessary with the calculator).

Programs that we prepared for use by the students included:

CDF and inverse CDF for normal, exponential, uniform/random samples for normal, exponential, uniform/random calculation of various sample attributes such as mean, median, range, mean absolute deviation for grouped and ungrouped samples,

significance tests and CI's for μ and σ^2 in normal samples;
significance tests and CI's for the exponential parameter;
significance tests about Bernoulli parameters;
various two sample tests of significance and CI's;
one and two-way analysis of variance (balanced design);
Scheffé contrasts
regression for one and two variables, including CI's for coefficient parameters, CI for a regression value, Prediction intervals, tests of significance for parameters;
Bayes' formula;
simulation of samples for certain discrete distributions;
expected values for functions of discrete random variables.

RESULTS

Generally, the students were enthusiastic about this approach. The calculator based statistics course became a popular course in the Management Science curriculum, in sharp contrast with earlier non-calculator approaches. A few students from each class actually began to choose Master's thesis projects based on the use of statistical analyses. Instructors of some follow-on courses (such as introduction to operations research and economics) were forced to revise their traditional approaches because the students had by that time developed considerable skill in preparing and using programs for their calculators. After an initial experimentation period, the Management Science Faculty endorsed this approach, and considerable efforts are being made to adapt virtually the entire Management Science curriculum to this approach.

There continues to be some resistance to these changes by some faculty members, however. A common criticism which I believe has some merit is that this approach makes the student dependent on a "mysterious black box". "What would one of these students be able to do if he or she were separated from the calculator?", I have been asked. An attempt to answer this brings us back to the idea of trying to decide what the students should be able to do as a result of completing their statistics course. I believe the calculator approach leaves more time for discussing the applications of statistics than did the traditional approach. I think that as a result they are better able to understand and utilize the results of statistical studies performed by statisticians in their organisations. And I can counter with the challenge, "What were they able to do after completing the traditional course?".

A second criticism I sometimes hear is, "Why don't you just run the analyses off on the computer and hand copies of the computer print-outs to the students?". We have experimented with such an approach, but without much success. First, it seems difficult to sustain motivation of non-technical students for such an approach. Second, there is merit, I believe, in having students handle data. This seems to give them a more accurate feeling that data contain information, that these are the basis of their inferences (rather than preconception or opinion), and that obtaining and handling good, clean data is not a trivial task. I have found the students seem to almost enjoy working with the data. Thus, it appears to me that looking at data sets and entering data into programs is, at least for small data sets, an asset in this approach. Some other potential criticisms to this approach were cited above. These include cost, mechanical reliability, and grading procedure issues.

CONCLUSIONS

The availability of low priced computing resources opens the possibility of restructuring traditional statistics courses. At the "low end", service statistics courses for non-technical students can take advantage of the capabilities of hand held programmable calculators. Such calculators have advantages (over use of computer terminals, for example) in terms of cost and portability. We are now at a point where students might be expected to purchase their own calculators for use in the classroom and at home. Using pre-recorded programs, a statistics course requiring virtually no algebra or computational skills can be designed for non-technical students. There seems to be some merit in having such students "handle" data, however. In such a course, emphasis is on the foundations and assumptions for statistical inferences, and in interpretation of the results of statistical analyses.

SIMULATION OF GAMES STUDIED IN THE SEVENTEENTH AND EIGHTEENTH CENTURIES

J Badrikian, IUT de Clermont-Ferrand

In the 17th and 18th centuries, most mathematicians took random situations from their social environment as subjects for their studies : particularly, games of chance which were played at that time.

We can see how probabilistic thinking developed from these props, by intuition. What Pascal called "biais". A reasoning which was, moreover, presented as "an Art".

To begin with, we suggest a statistical approach to these problems.

One can really understand a game if one often plays or watches others playing. The rules are then assimilated, the prognostics are given and tactics proposed. Practice will lead to familiarization with a random situation.

This is where the computer comes in : the games are simulated and the pupil can follow their development on the screen. The choice of representation (graphic, pictorial) is important. One must, at the same time, give all information on the development of the game in an energetic way (just giving simple score results would be very boring) with precision and clarity.

With the "Jeu de RECONTRE", *n* cards, numbered 1 to *n* are drawn successively from a bag. The player who withdraws a card bearing the same number as the order of the draw, at least once in the game, wins.

Representation on screen for
9 cards numbered 1 to 9

									Explanation
$\boxed{1}$	$\boxed{2}$	$\boxed{3}$	$\boxed{4}$	$\boxed{5}$	$\boxed{6}$	$\boxed{7}$	$\boxed{8}$	$\boxed{9}$	← Order of draws
7	1	6	$\boxed{4}$	9	2	3	$\boxed{8}$	5	← The 9 numbers are written down as they are drawn. "Invariants" 4 and 8 flicker: number (respectively 8) has been drawn on the fourth (respectively 8th) draw.

Simulation of a game is obviously quicker than the real game itself. The computer will therefore allow one to save time. As many games as one likes can also be played in a short time.

Secondly, the computer's function is to provide statistics. All results are memorized, the frequency calculations are done automatically and given at the end of the operation. The operator can do as many runs as he wishes. He can also develop sequences to any length.

Statistical results reply to questions already resolved elsewhere by mathematical reasoning and this underlines the bond between statistics and probabilities.

Here, the computer is only an aid to the pupil (he saves time, and he is discharged from doing frequency calculations) and it also helps him to maintain active interest in several ways:

- over and above the choice of the number of runs he can:
    - change the parametric values, for example, if two partners play with one probability of winning (0.5 for each player) he can modify this "chance" to his liking and examine the consequences on the different stages of the game;
    - moreover, having statisical replies, the pupil may wish to know how to resolve the problem by deduction.

Now would be the time to do a study, using probability calculus. And one may then need to use the computer again for the actual calculation of the established formulas.

It was Jacques Bernoulli who drew up the tables giving the winning probability for each stage of the game called "JEU DE PAUME". It took him a certain time to do this. A simple program can give them in a few seconds.

Today, I would like to present four simulations, realized on the screen of an APPLE II micro-computer which is connected to a printed output - which will eventually provide numerical results (frequencies, tables).

1. Simulation of the "Jeu de Paume"

This game was the forerunner of tennis. The "problem" is the same - although more complicated - as Pascal's game which the French call "Probleme des Partis" and the English call "Problem of Points".

The simulation is made for one chosen value of the parameter:

- number of games
- number of sets
- possibility of winning for each player.

Initial state (scores already reached) and the number of repetitions done from this state are left to choice.

The user can then modify the possibility of winning each game. Extra time is envisaged for three or more players.

On the other hand, after a certain number of repetitions we give the observer no possibility of following the game but give him just the final result and the frequencies.

2. Simulation of the "Jeu de Rencontre"

This game was studied by MONTMORT and was given the name of GAME OF THIRTEEN, because it consisted of 13 cards, numbered 1 to 13. It was taken up and discussed by DE MOIVRE, EULER, LAPLACE...

One could be interested in one or several "invariants".

3. Game of "The Three Urns" (Daniel BERNOULLI)

Three urns, each one possessing n balls of a given colour are used. We remove one ball from each urn. The ball from urn 1 is put into urn 2, the ball from urn 2 into urn 3 and the ball from urn 3 into urn 1.

From the development of the contents of the urns (which we see carried out on the screen) one can ask oneself numerous questions:

- contents of urns,
- time lapse for return to initial state,
- state of balance, ...

4. Game of Lottery

A lottery consists of n numbers. A draw consists of taking r numbers. After s draws we estimate the chance of only m numbers already given. (This is a question we often ask in France concerning our "JEU DE LOTO").

This problem was also studied by numerous mathematicians (EULER for example) and generalised the problem of throwing the dice by MOIVRE.

On the other hand, it illustrates the situation of drawing without replacement (the r numbers drawn) and draws with replacement (the s repetitions).

TOPIC 8: TEACHING STATISTICS WITH THE HELP OF CASE STUDIES

ORGANISER: PROFESSOR B EPSTEIN

CASE STUDIES FOR AN ADEQUATE UNDERSTANDING AND INTERPRETATION OF RESULTS BY STATISTICAL INFERENCE

M Borovcnik, University of Klagenfurt

NEED FOR NEW TEACHING STRATEGIES IN STATISTICS - PRELIMINARY CONCEPTIONS OF OUR UNDERSTANDING OF STATISTICS

In teaching probability theory some efforts have been made to support the understanding of the term probability as such. The foundation is laid by Kolmogorov's axioms (or some reduced system of axioms) but the intuitive way of teaching is by means of (Laplacean equiprobability and) relative frequencies. One very important aspect of relative frequencies is the empirical law of large numbers. For visualizing this empirical fact the idea of simulation has been (and still is) of great importance in teaching probability.(*)

The character of probability changes, if the field of application is widened to judgement of probability models itself. This is why one cannot rely on what has been developed within (a more or less mathematically oriented) probability theory. This bears the danger of an unfeasible transfer of a somewhat objective character of probability to statistics.

From the point of view of application it is statistical methods that count. Statistical ideas have penetrated modern sciences and ways of thinking. Fallacies, errors, and manipulation are widespread and characterise a helplessness in connection with statistical methods and results.

I want to characterize two ideal positions to tackle this problem in education at all levels:

(I) Probabilistic considerations dominate teaching stochastics, the mathematical level aimed at is too high, the connection to applications is abandoned, statistical problems are of minor importance.

This position has its origin in the following attitude: It is often thought desirable to work out the foundations, and to reason with logical preciseness, to trace back anything to the foundations. For stochastics this means that giving a complete representation of the theory of probability is considered the best way of teaching it. But theoretical connections are endless.

The mathematics of probability theory has many disadvantages:

- it precludes many people from occupying themselves with stochastics
- it obstructs the possibility of advancing to the relevant part of it, i.e. to statistics

---

(*) Cf. e.g. Heitele who apostrophizes it as a fundamental idea of probability

- it obscures ideas and concepts behind statistical reasoning
- it neglects primary stochastic intuitions, it fails to weaken false conceptions, it is not suitable to build the basis for stochastic schemes

(II) The other extreme position is that of teaching cookery recipes. Probability is too difficult, therefore it is reduced to a minimum, the remainder is an unrelated chain of methods trained by means of sterile examples - unconnected senseless mathematics.

Both of those extreme positions mislead us to apply statistical methods routinely (or else: to exclude this possibility of tackling problem situations) without adequately grasping the problem situation, without interpreting the results critically.

My thesis is the following:
Both positions are not apt to establish a picture of the real interrelations, they tend to give a biased objective conception of the possibilities of statistics.

There is a trend to dodge difficult mathematics, the old idea of simulation is extended to statistical questions in order to give statistics more weight and a better chance for being understood (Cf. e.g. Schrage). Decision rules are evaluated empirically, the whole test (estimation) problem is simulated under the conditions of a specified model and theoretical terms like risk of an error of type I or II can be witnessed in their long-run consequences. This performance should enable us to lay the basis for an operative understanding of central ideas and terms.

One of the central ideas is the following: comparing, evaluating, and grading a single value of a (test) variable within all possible values on the basis of an hypothesis.

This can be realised by theoretical reasoning (deriving the distribution of the test statistic) or by creating the basis for the comparison by simulation of the test situation (and empirically deriving this distribution).(*)

The concept of simulation may also be used to advantage in estimation problems: Visualising the long-run consequences of models and decisions (i.e. estimations) can give a better insight into notions and terms like unbiasedness, variance of estimation functions or confidence coefficient of interval estimates.

However, there is some danger of performing an unfeasible transfer of the probabilistic side to statistics:

(*) An idea going back to W.S. Gosset (in deriving the t-distribution he filled the gap with simulation studies) which was rediscovered for didactics only recently.

(1) One is tempted to consider the simulated data as real data, to interpret empirical meaningfulness into them and the terms which are visualised in their manifestations, to overlook that such simulated data are hypothetical data as they are gained on a basis of hypotheses. This is even complicated by the relation between the theoretical law of large numbers and the empirical one which has quite a different quality.

(2) There is the danger that one cannot clarify the conditions under which this simulation is done. In the probabilistic standard situation one is confronted with iid (independent identically distributed) realisations of a variable - transferred to the statistical standard situation, this would mean that one has iid-test-situations. There seems to be a very different quality of adequateness of iid-assumptions in some urn (lottery) situation and in the statistical test situation where one can hardly think of the realised situation as one of an iid-sequence of test situations (this fact should be worked out by case studies). Simulation of test situations alone does not meet the problem of teaching statistics in an intellectually fair way; it seems to oversimplify the problem.

(3) Difficulties in learning statistics are reduced to technical ones. It is agreed that there are special ways of reasoning in statistics: one has to grasp the various difficult technical terms as e.g. of the test theory. There seem to be no problems beyond that. This is by no means the case. A technical view on statistics increases the objective character of it and its results inadequately.

I agree that an operative understanding of central statistical terms and concepts is valuable - just as urn situations (and the simulation of them) are of value to visualise the probabilistic mechanism. However the field of real application is limited. The built-up understanding of the terms is biased thus impeding full understanding of the meaning, possibilities, and limitations of statistical reasoning.

Computers will soon aid us in calculating the necessary results. What we have to support by teaching statistics is to confirm our competence in facing seemingly objective results by means of statistics, to demonstrate the limitations and subjectivity of any statistical argument. It does not merely suffice to teach statistics, the conception of it which we facilitate is relevant.

- What are possible interpretations of statistically aided results? Difficulties with an adequate understanding are to be faced.
- What is the role of probability theory in understanding statistics?

Besides the details of the way of teaching statistics one has to avoid a biased understanding of methods and terms. We do believe that knowledge about statistical methods - meta-knowledge - will facilitate a better understanding. Furthermore, we think that case studies will be suitable to teach statistics as we intend to. We shall discuss some "meta"-case-studies to support our opinion and simultaneously give a picture of some aspects of our understanding of statistics.

Statistical problem situations are open problem situations. There is a much better chance to make this clear by case studies than by "conventional" teaching strategies:

One has to face problems like:

- judging information as to whether it is relevant to the problem at hand, whether it has its origin in qualified knowledge of the subject or is only something like a prejudice,
- deciding whether it is desirable to give an "overall"-treatment of the problem (e.g. the whole distribution in question) with no specified questions, combined with a low level of so-called background knowledge which may be false, in order to ask for details (concerning a parameter),
- deciding to include or to exclude various hypotheses, as it is often done only implicitly, by choosing a test statistic (e.g. χ^2 test or Kolmogorov-Smirnow-test),
- deciding which hypothesis is to be formulated as null hypothesis,
- judging random variability that is possible at a specified (given) sample size: significance in the technical-statistical sense does not mean the same as being of practical importance, failure to prove significance may be due to a small size sample only,
- arbitrariness in choice of technical procedures like different approximation rules, different groupings (χ^2-test) etc.,
- knowing and judging the peculiarities, possibilities, and limitations of special methods.

Knowledge of the problem and the field of application, information and competence to evaluate this information is most decisive in structuring a problem situation. By case studies we want to support our understanding of statisics:

As far as subjective components play a part (consciously or not) in the process of structuring a problem situation, in the choice of a test statistic, the results are of the same subjective character in spite of the mathematical methods used and the numerical exactness of the result.

The necessary competence is to be seen in two different modes:

(I) Knowledge in the field of application and the problem at hand

(II) Knowledge in statistics, in the peculiarities and limitations of the various statistical methods proper.

It is far more a convention that statistical methods are regarded as a useful tool in handling problems in various fields of application and in empirically oriented sciences. We have tried to outline why we think "conventional" teaching statistics gives a biased picture of this "art of reasoning", we have given some aspects of statistics that we consider especially important, now we are to proceed to discuss some "meta"-case-studies.

DISCUSSION OF THE CASE STUDIES

Case study 1:(*)
Number of boys in completed(**) families with four children.

Boys	0	1	2	3	4
Nr. of fam.	1	17	45	31	6

It should be investigated whether the data are consistent with the binomial model or if there are significant deviations from it (at the 5% - level).

We shall demonstrate various possibilities to tackle the problem:

(a) Information from population statistics is used to equate $p = 0.51$ - χ^2-test gives a prob value of 8%.

(b) χ^2-test with $\hat{p}$ estimated from the sample gives a prob value of 14%.

Neither χ^2-test indicates significant discrepancies between data and binomial hypothesis.

(c) 95%-confidence interval for the parameter p yields (0.508; 0.612). This result may be interpreted as follows:

- sample size too small, no judgement of a detail (of the distribution) seems possible
- information used in case (a) does not seem to be adequate, the percentage of boys may not equal that for the whole population.

(d) Variance test: $\frac{\text{empirical variance}}{\text{var. of spec. binom. model}}$ indicates a significant deviation - prob value 3%.

(*) No details of calculations will be given, only those aspects of results will be quoted which are needed to understand the line of discussion.

(**) i.e. families where no further children are to be expected.

(e) Confidence interval for standard deviation (equivalent to variance test): the theoretical standard deviation is in the upper rejection region of the connected test - the empirical variability being too small for binomial model.

In short, the results are as follows:

- χ^2-test fails to reject binomial hypothesis,
- confidence interval for the percentage of boys indicates: sample size being too small to judge a detail of distribution (local question),
- confidence interval for the standard deviation yields rejection of binomial hypothesis: empirical variability being significantly too small for binomial distribution.

This is an artificial and very academic case study, yet serving our purposes: to show that statistical problems are by no means straightforward. The different results do not cause logical difficulties but are somewhat disturbing if one wants to confer statistical results to reality. In the various ways of solution, different problems have been tackled.

χ^2-test tackles a global question: Is the number of boys binomially distributed or not? Any alternative is included in the consideration, the χ^2-test-statistic reacting to any deviation from the (binomial) model. Background knowledge is used only in the sense that one has a random sample. Other hypotheses are not referred to in case (b). In case (a) additional information on the percentage of boys is used but this does not seem to be justified in the light of case (c). -Using information one can make mistakes.

Confidence intervals for central tendency or for variability deal with local questions: a detail of the distribution in question is discussed. Treating local questions is not generally better than treating global ones: What is a possible interpretation of "no significant results" in cases (c) and (e)? Surely not that there are no significant deviations from binomial model, but only that there are no significant deviations concerning central tendency or variability. Only special alternatives are included in the consideration, for the test statistic is sensitive only in deviations concerning these details.

Using information to formulate local questions may be misleading, for using this information may be unjustified or simply incorrect. In test situations there is always the problem of judging information as to whether it is relevant to put concrete questions in order to get more information out of data - namely concrete answers:

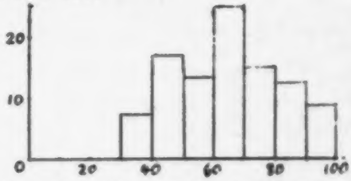
In complete families of four children the variability of the number of boys is reduced in comparison to the binomial model, for those families with four children having only boys or only girls tend to drop out of the population by getting another child.

One cannot escape the problem of local and global questions by arguing that χ^2-test is simply inferior in detecting deviations from a model because it uses only a small part of information contained in the sample. Another example could be given to show that this is not true. It depends on the view of the problem what alternatives are to be considered or to be excluded. There are questions in which any alternative has to be taken into consideration.

Case study 2:
Customers in a newly opened shopping city are interviewed at random. Among other questions they are asked to judge their contentment with the facilities of the new shopping city on a scale from 0 to 100.

The result of a sample of 80 persons is depicted in the histogram.

For further considerations it might be of value to replace this frequency distribution by a theoretical distribution. Is a normal distribution appropriate to describe the distribution?



(a) Test for normality: χ^2 yields a prob value of 34% - normality seems to be quite appropriate.

(b) Test for uniformity: χ^2 yields a prob value of 8% - one cannot exclude uniformity (at the 5% - level).

(c) Let us consider increasing the sample size to 94, the percentages being unaffected.
Normality: p. v. 3%; uniformity: p. v. 4%.

By varying the hypothesis being tested or by varying the sample size one can see: the choice of the null hypothesis is very critical - one must judge the sample size for relevance. The results are very sensitive to these "parameters".

Case study 3:
A firm is interested in improving prevention of accidents. For that reason one wants to investigate (among other things) if there is a significant deviation from a uniform distribution of risk over the days of the week.

Mon	Tue	Wed	Thu	Fri
26	17	22	24	31

(a) χ^2-test for uniformity yields a prob value of 50%. The data cannot at all be regarded as unusual - no rejection of uniformity hypothesis.

(b) Separate 95% -confidence intervals for the percentage of accidents occurring on each day of the week have a width of approximately 15 percentage points, e.g. for Monday (0.14: 0.29). The separately calculated confidence intervals are very wide
- sample size seems to allow no conclusion other than its being too small (simultaneous confidence interval would be even wider).

(c) Supposing there is past information that accidents on Mondays and Fridays are more frequent. Test of $p = 0.4$ against the alternative $p > 0.4$ would just reject the uniformity hypothesis relative to Mondays and Fridays (at the 5% -level).

In cases (a) and (b) the method used was varied - again a choice being open between local and global aspects of the problem. If it were of practical importance then one would decide to get more data.

In case (c) another aspect is considered: The problem of whether one can use information available in order to put more precise questions to get more information out of data. What if this information is just a cursory impression, or even just a prejudice? What if it is extracted from the same data by means of which it should be tested? Judging this information for relevance to one's problem is clearly subjective. The same subjective feature lies in the variability in case study 2 and in the choice of the method in cases (a) and (b).

STRUCTURING THE PROBLEM SITUATIONS - CONSEQUENCES FOR THE CONCEPTION OF STATISTICS

(i) Choice of test statistic

By varying the choice of methods, choice of null hypothesis and information about the problem at hand we have seen that in any statistical (test) situation there is a large spectrum of possibilities. By choosing a test statistic we implicitly or explicitly discern two classes of deviations:

- important deviations, the test statistic being sensitive to them,
- deviations of no importance, the test statistic being insensitive to corresponding alternatives.

(ii) Local and global questions

There are very different aspects connected with these opposite ways of questioning. Does it really help if χ^2-test leads to rejection of the binomial model (case study 1)? What if one cannot reject the binomial hypothesis by the variance test?

One has to balance advantages and disadvantages of all the possible results and their interpretation under rival ways of tackling the problem.

(iii) Information

Justified use of information may lead to concrete (local) questions and in consequence to concrete answers which do yield much more information about the problem at hand (e.g. direction of deviations). On the other hand use of information means exploiting background knowledge which may be false.

One has to balance advantages and disadvantages of using information: Is it better to mark out special alternatives to get more information out of the data thus risking making an error or is it better to avoid this risk and include any alternative?

(iv) Relevance of sample size

Statistical significance is not equivalent to practical importance. All results are relative to sample size, so there is a need for judging the magnitude of random deviations to be tolerated at a specified sample size by direct methods.

- In case of no rejection one has to ensure that this is not due to a small sample size, a sample size at which random variation is so large that even considerable discrepancies between data and model have only a small chance of being detected.
- In case of rejection one has to evaluate magnitude and relevance of the deviations by other (direct) methods than just by saying that e.g. it exceeds the 95% -fractile of the test distribution.

Such judgements and the coherent acts of choice - including, excluding alternatives in the consideration; balancing advantages and disadvantages of local and global questions; evaluating information; judging sample size for relevance etc. - should not be made arbitrarily or simply implicitly without deliberating over their consequences.

Expert knowledge in the field of application, the problem situation at hand as well as the peculiarities of statistical methods in question should form a basis for structuring the problem finally leading to the model situation of mathematical statistics or, if simplified, the corresponding simulation situation.

Of course this structuring process is markedly subjective, many problems being coherent in it: Isn't even using information in order to structure one's problem manipulation? It depends on one's point of view how information is evaluated.

One consequence is that results gained by statistical methods are inseparably connected to the whole context of the process of problem structuring and all its subjective components. Without communicating these context conditions statistical results are nothing but bare bones: void of any useful interpretation, or: capable of any wanted interpretation.

MEANING OF STATISTICAL RESULTS FOR REALITY

By varying the conditions the case studies make clear that statistical methods are very doubtful if their results are to be interpreted adequately. Regarding the complexity of problems in connection with these judgements something like a return to a "weakened form of determinism" could be desired: To have only one test statistic to check discrepancies between data and model would give the results at least a normative character. The whole process of structuring with all its subjective components would be eliminated. The case studies, however, show that there are interesting questions which could not be answered in this way. Thus, let us forget this idea.

What do statistical results really mean in the light of the discussion of the case studies?

What does it mean that the test statistic XX has realized a value of c, the significance of the result being beyond the 5% -level? Other test statistics would have yielded a different significance level.

It is a usual convention that risk of a type I error is regarded as property of the test procedure in question and not of the tested hypothesis. The testprocedure will reject a correct null hypothesis only in $100\alpha\%$ of performances of the test - α being also called the size of the procedure.

Besides the general difficulties in transferring a meaning of this long-run property to a singular test situation there are many aspects to be considered:

- One includes only error of type I, errors of type II depend on the alternatives determined by the structuring process.
- How can risks of a type I error be compared with each other in different test situations? One does not always test correct hypotheses, the tested hypotheses not always being of the same importance to reality. For example, simulating test situations to give an impression of the corresponding notions misleads someone to interpret empirical meaningfulness into such terms.
- A frequency interpretation of the size of a test is restricted to "clinical" iid-conditions thus being a model conception - the case studies with their structuring process indicate that this could only be an hypothetical (urn) conception for the problem at hand.

- The process of problem structuring makes use of information and other subjective components - the resulting discrepancy between the outgoing problem situation and the model situation being unique in any case. The distinct contents of background knowledge used in the structuring process makes statisical results uncomparable because they lose their sense without the whole context.
- Comparing and judging a realised value of a test statistic with a theoretical distribution is clearly a model comparison. Simulating this process cannot wipe out these theoretical components. Besides it may lead to false impressions because the iid-conditions are clearly satisfied in the urn (simulation) situation but by no means in the real problem situation.

All this means that the singular test situation is compared to a sequence of independent identical test situations. There is one singular realisation of the test situation - with all the subjective components of the structuring process - the real problem situation being by no means identical with the model situation in question. Nevertheless the model situation is thought to be identically repeatable and the result of this hypothetical process is used to judge the initial real problem.

The discussion of the case studies stresses that statistical results can only be interpreted hypothetically. The concept of iid-realizations of a random experiment - e.g. playing roulette - is uncritically transferred to statistical reasoning. One loses sight of the fact that it is the process of problem structuring that leads to the "clinical" standard situation. One "creates reality" without reflecting this.

My thesis is the following:
The possibility of interpretation in the clinical iid-situation of probability theory is extended to statistical reasoning by an unfeasible transfer leading to a false objective conception of statistics itself.

The notion of probability seems to have changed its character from objective to subjective via the extended field of application from probability to statistical reasoning. It is desirable to investigate whether it would be of advantage to understand the notion of probability right from the beginning the way the subjectivists do. Statistical methods do not gain objectivity by the fact that many people acknowledge their merits and adhere to the usual mode of interpretation - statistical results being valuable only in connection with a balanced view of the problem situations which ought to be determined by expert knowledge in the field of application and statistical methods - they are not to be separated from the whole problem context and its structuring.

One always has to point out the discrepancy between the initial problem situation and the model situation. Statistical methods cannot replace knowledge of the problem at hand - statistics per se cannot give insight - it can only help finding one's way through.

CONCLUDING REMARKS

The discussion of the case studies and the continuing discussion of the problems of adequate interpretation of statistical results should provide

- a picture of our understanding of this "art of reasoning",
- sufficient support to show that case studies(*) offer an excellent opportunity to facilitate a better insight into the problems of mathematising real situations and interpreting the results adequately.

In this light statistical problems appear to be open problems.

Several - subjective - choices and judgements have to be made to establish the model situation of mathematical statistics:

(i) Choice of the test statistic

(ii) Choice between local and global questions

(iii) Judgement of information for relevance

(iv) Judgement of relevance of sample size

Case studies should give a chance to see that statistical results can only be interpreted hypothetically - their relative simplicity not only being artificial but also being a basis to facilitate an understanding of this feature of statistics. One cannot interpret statistical results adequately separated from the whole problem context including its structuring. The shortcut "the h_o will not be rejected at the 5% -level", where you can take any h_o and where the problems are exchangeable, is misleading and very academic.

Another step is understanding statistical concepts and the meaning of statistical results can be achieved by doing statistics in an exemplary way, i.e. by means of project studies. We have not tried to tackle the problem of project studies simply because we lack experience. Nevertheless we think project studies should permit the insight that personal purposes and intentions do affect this process of problem structuring and thereby final results if this is reflected in a specific phase. Maybe the problems of real application become too powerful so that we think case studies could facilitate this conception of statistics more easily.

(*) For further case studies especially dealing with limitations and possibilities of statistical methods see [1] - where, for example, a parallel between X^2-test for independence and the Pearson correlation coefficient is outlined.

On the whole, case studies (and exemplary project studies) can permit the perception that not only knowledge of statistics (in the mathematical-technical sense) but especially expert knowledge in the field of application, of the problem at hand, and some "meta"-knowledge about possibilities and limitations of statistical methods in question will and should play a decisive role to compete in situations with uncertainty. The intuitive way to teach statistics should be by means of case studies. Case studies should have the same place in teaching statistics as simulation (*) has in elementary probability.

(*) In connection with demonstrating the consequences of certain probability models - not in the sense of Monte-Carlo studies.

References

[1] Borovcnik, M. (1981) Eine Analyse des χ^2-Unabhängigkeitstests - Möglichkeiten und Grenzen, Schriftenreihe Didaktik der Mathematik, 3, p.35-48.

[2] Heitele, D. (1976) Didaktische Ansätze zum Stochastikunterricht in Grundschule und Förderstufe, phil. diss., Dortmund, 1976.

[3] Schrage, G. (1981) Entscheiden und Begründen - Leitlinien für den Statistikunterricht, Schriftenreihe Didaktik der Mathematik, 3, p.86-101.

IS A SCIENTIFIC LEGEND A COMPREHENSIVE CASE STUDY?

V V Shvyrkov and A C Davis, Rider College, Lawrenceville

We were standing in the back room of historic Bainbridge House in Princeton, New Jersey contemplating this saying of the late Albert Einstein:

"There was this huge world out there, independent of us human beings and standing before us like a great, eternal riddle, at least partly accessible to our inspection and thought. The contemplation of that world beckoned like a liberation".

A child next to us asked his father, "What's Einstein? Sounds like a monster". Father: "Some people think so. He was an inventor". Child: "What did he invent, some kind of telephone?".

We are reminded by this nine-year old and his father of our students taking required courses in business statistics. They seem to encounter much difficulty compared with other business courses they take. It is possible that much of the difficulty arises because students perceive statistics as a bewildering array of arcane techniques, bristling with hocus-pocus formulas, the eternal mystery of degrees of freedom, brain-teasing word problems, and a Rubik's cube model of the real world, with data grouped and scrambled beyond hope of reconstruction.

As a result of students perceiving statistics as an overabundance of seemingly unconnected methods and problems, statistics, rather than being viewed as a powerful research tool, is viewed simplistically as the application of a collection of many methods to hundreds of disjoint problems. At best, such an approach to the study of statistics is unrealistic. On the surface, students are being taught how to use statistics practically and with discretion. But their perceived satisfaction of knowledge exceeds reality, so that their little knowledge is a dangerous thing. Unfortunately, the simplified approach to teaching statistics presupposes without verification that all underlying mathematical and methodological assumptions are met. Finally, typical case studies are designed to illustrate only a limited number of concepts.

The authors maintain that much of the trouble we have mentioned could be eliminated if statistics were presented as a scientific investigation of a single comprehensive real-life case study, which we conditionally call a "scientific legend".

I. COMPREHENSIVE CASE STUDY ... PLUS

We have developed a method of teaching Introductory Statistics I and II at Rider College using this case-study approach. Why the name "scientific legend"? Well, a legend is a story more real than reality (distilling the essentials into a compact "case") and our legend is this essence pressed into the use of science! A story teaches reality by distillation, but is neither a small intricate puzzle nor an overwhelming mass of confusion.

Existing case studies on business statistics are based either on samples or on a population data set. Case studies based on samples are designed to present statistics as a complex of disjoint problems. On the other hand, case studies based on a population data set are designed to provide students with unified statistical problems rather than theory. Both types of case studies do not consider statistics as a science, nor do they teach students how to verify the underlying assumptions under which statistical methods are valid.

A scientific legend (SL) is a comprehensive case study based on a population data set. Our particular scientific legend consists of real-life data (150 family budgets) constructed from average values taken from the 1977 Statistical Abstract of the United States. Each budget contains values for nine family characteristics (variables): husband's education; husband's job position; husband's height, wife's height; son's age; family income; number of rooms; expenditures on housing; and, annual bread consumption.

Our broad goal in constructing the legend was to create a vehicle that could be used to present statistics as a scientific investigation of a connected set of real-life economic problems as opposed to a collection of disjoint questions. Thus, in our classes a single set of data is used throughout the semester to illustrate all principal problems and methods usually covered in a business statistics course.

At the beginning of the course, the legend is used to explain the theory of probability. This material is crucial as it forms the foundation for all statistical theory to be covered later on. With the help of the deductive method the legend is used to explain laws of generating samples from populations (in this case, the population consists of the 150 family budgets).

With this background knowledge, the legend is used as the basis for considering generalized statistical measures. Here, each student works with a unique subsample of observations (32 or more budgets) and studies the remainder of the topics. The scientific legend is designed to present statistics as a science through the following stages:

1. Analysis of statistical data set quality.
2. Verification of underlying assumptions.
3. Application of statistical methods.

Audio-visual aids and computer-assistance enable presentation of a scientific legend and its concept. Let us consider the main three conceptual stages of the scientific legend:

1. Success of any statistical research requires coordination between theory (the theory of probability) and practice (data analysis). The theory of probability is based on the foundational assumption that a population is homogeneous (e.g., an honest coin, an unloaded cubical die). Coordination between theory and practice requires the same foundational assumption for the statistical data analysis. Therefore, we can state that statistical data analysis should be based on a homogeneous statistical population.

Quality in economic research depends upon the generation of one-dimensional and n-dimensional homogeneous statistical distributions. These distributions can be considerud as a bridge between the theory of probability and statistical analysis in economics. Our student source population is, by design, homogeneous.

2. The verification of underlying assumptions is the second important stage in the use of a scientific legend. On the basis of SL we verify the following underlying assumptions of regression models:

   a. Independence between factors.
   b. Homoscedasticity of conditional variance.
   c. Normality of the residual variable.

For the solution of all these problems the Chi-Square test statistic and the K-S test statistic are employed. Classical theory provides a weak inference because it is based on a single sample subject to errors. We suggest that a stronger inference be constructed on the basis of multiple samples.

3. The scientific legend is designed as a vehicle to teach business statistics on the basis of a "realist" approach. The realist approach follows these rules:

   a. Difficulties in statistics should be detected and explained, not avoided.
   b. Practical explanation should be made first, and abstract theory second.
   c. One should teach "Business Statistics" through practical exercises, not through riddles.
   d. In business statistics, linking of terms, definitions, and the concept of data quality (and quality, not necessarily quantity!) is very important, remembering Richard Hamming's dictum, "The purpose of computing is insight, not numbers".

e. Statistics should be taught on the basis of continuity among topics. We offer the following logical sequence in our courses:

Foundations of statistics
Generalised measures
Statistical analysis of regularities
Statistical inferences
Time series
Index numbers

II. AUDIO-VISUAL AIDS

The method of the scientific legend can be practically taught only with the help of audio-visual aids and computer assistance. The purpose in constructing the method was to provide a vehicle that can be used to present statistics as a scientific investigation of a connected set of economic problems (in contrast to a collection of disjoint problems). Through this method, students, working with unique subsamples, analyse a connected set of ungrouped data. Emphasis is put on the verification of statistical data quality before the application of statistical methods.

The traditional chalk-and-blackboard teaching technique, adequate for the customary simplified examples, is insufficient when one deals with ungrouped data. Therefore the realist approach requires transparencies for the instructor and a special student manual for the class. Since student problems have individual answers, computer assistance is required for the generation of homework and exam answers. A standard text provides theoretical material to reinforce the lectures, and videotapes or films provide a backup and review, but not a substitute for the lectures.

Statistics through transparencies is characterised by many illustrations in pictures, graphs, tables, and formulas. Transparencies have been elaborated with didactic attention to the following points:

a. Principle of the still: one idea in simple form with few words
b. Window method (progressive disclosure)
c. Link (sequence) method
d. Method of repetition

Statistics films are mainly based on materials prepared for transparencies. Designed primarily for students missing lectures or wishing to review, they are characterised by fewer words, many illustrations, and a quicker speed of presentation. Our experience has shown that the speed of presentation of materials in a film can be as much as four times that of an ordinary lecture. Since the films will be housed in the college library, they are easily accessible to students, and since they are all based on data contained in the "scientific legend", continuity in material coverage is more easily effected.

The method of the scientific legend lends itself to the development of interactive teaching and instant feedback on the correctness of homework through the use of terminal oriented mainframes or minicomputers. Students are therefore able to pace themselves, to gain reassurance, and to avoid series of errors due to interproblem dependencies.

Proposed Benefits

Sicne unique subsamples (consisting of 32 or more family budgets) can be drawn from the population, each student works with a unique set of data throughout the term for homework and, if the instructor desires, examination purposes. Although comprehensive case studies (or so-called "practice sets") have been successfully used in other business courses (e.g. accounting), the authors are unaware of any attempts to use this approach in the teaching of statistics. That the legend can be used to generate individualised comprehensive case studies is almost certainly novel. The use of adjusted real-life, rather than purely fabricated, data is designed both to enhance student interest and involvement and to present statistics as a powerful yet practical research tool.

Since the measurement level of the family budget data with which the students are working varies (nominal, ordinal, interval, ratio), it is easier to explain the circumstances under which particular statistical methods are appropriate or inappropriate. The goal of avoiding indiscriminate application of techniques without prior verification of mathematical assumptions is therefore continaully stressed. In our view, this constitutes a more realistic and beneficial teaching approach.

The use of individual samples from a common population generates the additional benefit that consciousness of variability in sample data has already been raised when the subject of sampling distributions comes up. The probabilistic nature of statistical inference is apparent in contrast to the "proofs" of other scientific and logical systems. The use of computers by students at a success and sophistication level far above that experienced in beginning programming courses should help to reinforce positive experiences in the use of these tools.

III. COMPUTER IMPLEMENTATION

Computer-assisted instruction has not taken over the world; certainly not in statistics. Perhaps it is that the existence of computer-aided instructional methods has preceded their essence (Michelson, 1970). We may, just as we need new wineskins to contain new wine, need innovative theory to contain new methods. The method of the scientific legend is an attempt to provide just such a theory most suitable for implementation on computer media.

"Experimental physics is the strongest argument for scientific realism. The current fad in philosophy is due to doing theory, not experimentation. Nothing but experiment is so lacking in science.... A good experimenter will bring forth phenomena that have not even been seen before." (Hacking, 1982). In statistical education, data and examples are particularly important in the impartation of insight. We only need continually to fight the tendency to lose the insight among all the numbers!

The SL method is so regular that we are able to implement it rather straightforwardly on the computer. [We are now using a Digital Equipment 2060 and writing in an extended Fortran]. Not that any system of this magnitude is trivial or small to implement, but its feasibility has been heartening. We are taking a structured approach to the programming, keeping computational, printing, data management, and interactive functions as separated as possible. Each student at the beginning of the course is assigned an id. number and a password. The id. numbers are the keys to the table of random numbers and the password is used to prevent one student from tampering with another's results or from seeing another's grades.

Several methods of handling the generation of samples might be possible. We could try to do this using published random number tables, and devise a method for each student to select suitable numbers within the range of family numbers in the population. Alternately, we could use the computer to generate random numbers. This is the method we are using. At present, we are taking advantage of a prime-modulus generator supplied with our Fortran function package, generating numbers in floating point form, and by a linear transformation mapping them onto integers in the range of 1 to 150, matching the population. We have tested the output of this generator and function for uniformity of the distribution and are satisfied that it provides numbers suitable for these educational purposes. Before converting the programs to another computer system, we will want to research a portable number generator that will provide numbers that do not vary with hardware and method of arithmetic. The new IEEE standards on computer arithmetic may eventually help in this matter. Of course, these "random" numbers must not be random with each computer system or execution, or the student id.s would be untraceable and meaningless. Since they are generated by a deterministic algorithm, they are repeatable.

We are preserving on disk whole sets of random numbers sufficient for several years' of students. These are indexed with the id. key number given to each student at the beginning of the term. Then the random numbers are used in the computer just as the student uses them: to select subsamples from the population, which is kept on another disk file.

At present, we have printed answers to all homework and examination problems. These can be used to check the students' final answers. For some chapters we also have generated detailed numerical solutions for each student which are useful in diagnosing where the student has gone astray in his work. We intend to finish these solutions and implement a further refinement in the next year, namely, a program to check a student's homework interactively at a terminal.

The interactive method will work as follows. Students log-in to a special account with execute-only privileges, giving their random number key id. and password. The defaults to this special account take over from there. The student is asked for which problem he wishes to check his answers. Answers are checked, telling the student only whether he is correct within a) a high degree of accuracy or b) within an acceptable degree of accuracy or c) not correct. The accuracy required is generally plus or minus a few percent of the correct value, but depends upon the particular problem and the sensitivity of the calculation. Students are not given the correct answers if wrong, but are sent back to try again until they have mastered the problem. The relative accuracy and the number of times attempted to get within the required accuracy is recorded in the grade table data area for this student. An average score is kept on a running basis according to a pre-announced formula for each student and is available at each session for his perusal. This feature allows an instant measure of progress, a satisfaction we find students crave. Periodically, the instructor may obtain summary printouts on student progress and can call attention to those students who have not progressed in a timely fashion. These reports also serve as backup documents on the grading in the outside event of disk failure or security compromise of the system. In this process virtually all grading labor has been automated, freeing the instructor for more profitable activities such as working individually with students.

A further level of computer implementation would involve computer-assisted instruction (CAI) wherein the student could be walked through the solutions to problems he is not able to get on his own. This would involve a far greater investment in terminal time than the checking of pre-worked answers, as well as a greater level of program sophistication. At present we plan to leave this function to the instructor who can walk through problems with the help of detailed printouts that can be prepared either ahead of time or on demand for each student.

Program Structures Required (see Fig.1)

The following programming structures are required:

1. An executive routine which handles interaction with the students at the terminals and calls the appropriate calculation subroutines.
2. A set of subroutines to implement each of the statistical methods to be taught (e.g. mean, dispersion, regression analysis, analysis of variance, etc.). These routines have interface with the data fields of the legend or legends.
3. A table containing data for the particular legend being employed.
4. A routine to generate tables of random numbers keyed to student number (once for the term).
5. A table of grades and a record of interaction with each student logging-in.
6. A record of statistical information about the operational experience of the system for later analysis (and, we hope, future papers!).
7. Routines for instructors' retrieval of grades and data as well as prepared student materials such as sets of random numbers, detailed problem solutions, exam blanks, etc.

Data and Program Security and Implications for Automatic Grading (see Fig.2)

A number of security problems arise when students are left alone to work on the computer. Our problem is how to minimize these exposures in a practical system. The primary security comes from the nature of the legend method itself. Since each student has individual assignments, no student can copy another's answers. For homework, this is probably sufficient, since the burden of doing assignments for more than one student is too great for anyone. For exams and quizzes other problems arise. One student could pay or coax another student to take the exam for him if it were to be directly entered into the computer without proctoring. The computer cannot see; any identification a student might be asked to provide at log-in could be transferred to his proxy. The following method is suggested to preclude this possibility while maintaining a high degree of automation.

An individual pair of identical answer sheets (possibly with carbon paper) are prepared by the computer system. Exams are administered by the instructor in the usual manner. One copy of the answer sheet is collected at the end of the period and the other is kept by the student for entry at the terminal as in the manner of the homework. After a sufficient time, the computer prints for the instructor copies of the answer sheet filled-in with the students' responses.

These can be spot-checked or candled with the retained copies for differences. the computer will add up scores as usual; any exceptions from the verification (retained) copies can be entered by the instructor. If an optical or mark-sense scanner is available, the student participation in entering answers is not necessary, as the exams can be entered directly in machine-readable form.

Of further consideration is the matter of tampering with the system. Data files (including student grades) and system programs must be protected from alteration by enterprising or malevolent students. Several measures can be taken. First, the programs can be stored in load module form only, which is virtually indecipherable; it will only execute and cannot be read. Second, data file names can be camouflaged and kept secret. The references to them from the programs will not be decipherable in the load modules. As a further security measure, the grade data can be encoded so that even if students were to stumble upon the files, they would be unprintable. Also, they would be unalterable without destruction. Additionally, the protection of the operating system should be used to isolate the vital files from all users other than the programs of the legend system with a "need to know". Finally, programs and data segments are kept in separate accounts, separately protected and with camouflaged account names. A further hazard can be created by the student who wishes to abort the program in mid-session or by system crashes. We therefore find it necessary to trap terminal interrupts (control-C on the DEC-20) so as not to allow the student to abort suddenly, and we must keep a record on disk of the partial progress of each student.

A Review of Our Experiences to Date

One of the ironies of the CAI experience is that computers can provide a level of "personal" attention to the student that faculty are unable to do for lack of time. This is particularly important in light of the large sections now encountered in business schools and other departments teaching "practical" subjects. Further, the workbook and the computer allow the fast students to proceed at their own pace and slow students to take their time. Some of the homework deadline pressure is relieved, since assignments do not have to be due on particular days, but students can enter their results any time computer terminals are working and available, which in most hours of the day and night. The homework is graded by formulas for accuracy and with the speed and patience that only a computer possesses. Tests and quizzes can be given with the frequency that an instructor desires, knowing both that grading tedium has been relieved and that all does not rest upon the exams, but that homework is also a reliable measure of student performance. As in all courses, students want frequent evaluation of their progress. This is particularly so in a grade-conscious generation. The use of computer terminals makes this feasible.

We believe from our experiences to date that this will have the further effect of keeping up morale and interest in the course.

From our experience of teaching by the method of the scientific legend in the last year, we offer these observations about what we have done right and what we wish to alter for next year: First, what is right: Students work hard and gain much experience in practical statistical calculations. Many students find an interest in this method where they had not from conventional approaches. For the teacher it is a relaxing and self-cueing method. (The authors do not find they need to bring any notes to class). Nevertheless, some students say they are addicted to chalk dust (author Davis personally is allergic to it!) and they cannot learn if there is not enough of it flying during a lecture. One student told another, "This is the most bizarre way of learning I've ever seen". However, we have found that after a few weeks the conservatism and the shock of new procedures wears off. (Our methods if observed resemble more the corporate and continuing-education methods that are now prevalent than they do the traditional academic classroom atmosphere).

At the beginning of the course term we would emphasize more attention to detail. We have found many students are sloppy in their habits and we feel one important lesson in a beginning statistics course is emphasis upon accuracy and form. Round-off is another subject that remains a mystery to students unless specifically stressed. Topics of sensitivity of the various methods to errors and the effect of roundoff, e.g., in regression analysis, will be emphasised more in the future. Students will be encouraged to develop a sense of judgement in these matters and to present final results with an appropriate degree of accuracy. Until a complete text is developed with this method, students will be encouraged not to miss class, or in that rare event to review by way of the videotapes. Since this is a unified method, students can lose depth by falling seriously behind. There is always a temptation in this method to work by cookbook, but serious errors in judgement will result thereby.

In order to break what can become after a long time "slide hypnosis" and to afford more opportunity for student interaction, we plan more frequent short quizzes and student-led seminars on the material. We wish to emphasize the importance of obtaining a good personal calculator with scientific functions, and possibly with programmability and elementary statistical operations built-in. We believe many students have been slowed down unnecessarily by tedium in using crude four-function calculators. The college is obtaining a graphics software package and a full-color plotter on its computer next year and we hope to connect this so that students can get plots of their individual sample distributions. We believe this will be another feature to increase student interest in the material.

CONCLUSION

Our motivation in developing a "scientific legend" has been to encourage students to view statistics as a unified set of procedures by which to scientifically test probabilistic observations about the real world. The existence in the last few years of certain technology on the market, viz., terminal-oriented computers and inexpensive hand-held calculators has made the contemplation of this approach possible. We are attempting to motivate them to work with real-world, ungrouped data, and have been developing a computer package designed solely for instructional purposes.

References

[1] Anderson, Donald T. and Rivers, Richard A. (1980) *Practice Problem in Financial Accounting: Saluki United Vending Service, Inc.* Houston, TX: Dame Publications.

[2] Cooley, William W. (1969) "Computer Assisted Instruction in Statistics". *Statistical Computation*, Roy C. Milton and John A. Nelder, eds., New York: Academic Press.

[3] Forsyte, Alan B. and Freed, Jas R., (1979) "An Evaluation of Computer-Aided Instruction in an Introductory Biostatistics Course". *J. Computer-Based Instruction*, February, 5, 57-62.

[4] Gallant, A. Ronald, (1973) "Some Arguments Against the Use of Statistical Packages in Teaching Statistical Methods". *Proc. Computer Science and Statistics: Seventh Annual Symposium on the Interface, October 18-19*, pp.223-225. [Plus, other papers this volume].

[5] Gross, Beatrice and Gross, Ronald. (1980) *A Review of Innovative Approaches to College Teaching*. Sarasota FL: American Accounting Association.

[6] Hacking, Ian (1982) (Stanford University Department of Philosophy), "Experimentation and Scientific Realism", Princeton University Philosophy Seminar, January 29th.

[7] Hobbs, Gerald R., Jr. "Some Advantages of Individually Generated Data Sets in Statistical Methodology Courses". - in forthcoming volume by Academic Press, pp.465-469.

[8] Michelson, S. (1970) "The Existential Reality of Educational Production Functions". Washington, DC: Office of Education.

[9] Shvyrkov, V.V. (1982) *Scientific Legend in Applied Probability and Statistics*, (Vol.I) Lawrenceville, NJ: G. Throwkoff Press (Student workbook).

[10] Shvyrkov, V.V. (1982) *Scientific Legend in Statistical Science*, (Vol.II) Lawrenceville, NJ: G. Throwkoff Press (Student workbook).

[11] Tobias, S. (1969) "Effect of Creativity, Response Mode, and Subject Matter Familiarity on Achievement from Programmed Instruction". *J. Educational Psychology*, 60, 453-460.

[For further references on computer-aided instruction, see our paper (A.C. Davis and V.V. Shvyrkov, "Computer Assistance in Teaching Statistics Using Method of the Scientific Legend") Proc. 1981 Joint Statistical Meetings, (August 10-13, 1981, Detroit MI.)].

BEST COPY AVAILABLE

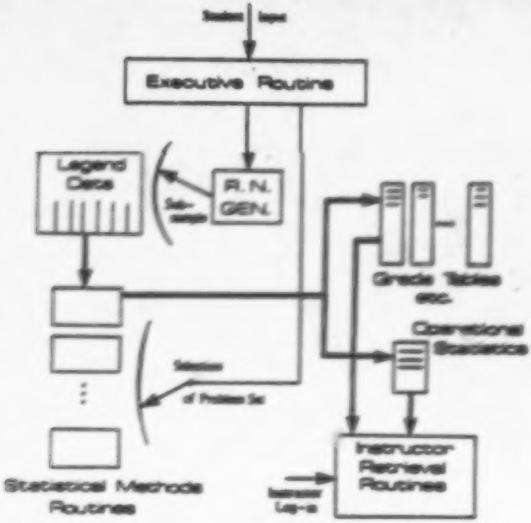
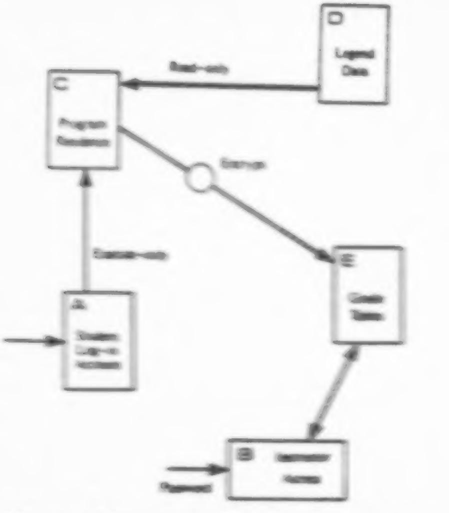


Fig. 1 Program Structures



Squares Represent File Protection Areas

Fig. 2 Program & Data Security

FAMILY BUDGETS, 1977

#	HUSBAND			WIFE'S	SON'S	INCOME ($1000S)	NUMBER OF ROOMS	EXPENDITURE ON HOUSING ($1000S)	CONSUM. OF BREAD (IN POUNDS)
	EDUCATION	JOB POSITION	HEIGHT IN INCHES	HEIGHT IN INCHES	AGE (IN YEARS)				
1	H	S	71.0	71.0	12.9	33.0	6	6.83	315.0
2	E	N	67.0	59.0	7.4	5.0	2	1.12	302.5
3	E	S	69.7	69.0	15.9	33.0	6	6.50	325.0
4	E	N	64.3	62.1	9.2	8.8	2	2.86	305.0
5	C	S	71.2	69.5	12.9		[illegible]	7.50	325.0
6	H	N	66.1	[illegible]	[illegible]	15.3	[illegible]		
140	[illegible]	[illegible]	70.9	65.4	11.8	14.1	4	3.16	321.0
141	E	N	70.1	65.2	11.7	13.8	3	3.42	319.5
142	H	N	66.9	66.5	9.3	15.2	3	3.47	281.7
143	H	N	71.4	63.7	8.0	10.8	4	2.62	297.7
144	E	N	69.4	65.1	11.6	13.5	3	3.31	318.7
145	E	N	68.2	69.0	13.8	19.5	3	4.50	302.5
146	E	N	69.7	67.2	12.9	17.0	4	3.91	329.0
147	H	S	72.3	67.8	12.6	23.7	3	4.90	285.0
148	E	N	69.9	63.2	10.3	10.8	4	2.37	313.7
149	C	S	70.6	63.4	9.4	13.3	3	3.25	305.0
150	H	N	70.4	63.4	8.0	10.5	4	2.12	296.8

EXPLANATION OF SYMBOLS

E	ELEMENTARY SCHOOL EDUCATION	N	NON-SUPERVISORY POSITION
H	HIGH SCHOOL EDUCATION	S	SUPERVISORY POSITION
C	COLLEGE EDUCATION		

Fig.3 The Source Population

```
 EXECUTE CHAP2
LINK:  Loading
[LNKXCT CHAP2 execution]

 PROGRAM TO PRINT ANSWERS TO PROBLEMS IN CHAPTER  2

ON DISK ARE LEGENDS A THROUGH E, EACH CONSISTING OF
160 STUDENT CASES OF 32 RANDOM NUMBERS EACH

TYPE IN LEGEND LETTER (LIMITS: A TO E) B
TYPE IN STARTING AND FINISHING STUDENT NUMBERS,
SEPARATED BY A BLANK  17 17
TYPE IN STARTING AND FINISHING PROBLEM NUMBERS,
SEPARATED BY A BLANK  18 21
FOR REPORT IN STUDENT-MAJOR ORDER, TYPE STUDENT ELSE TYPE PROBLEM
FOR PRINTED OUTPUT, TYPE THE WORD print AND PRESS RETURN;
FOR TYPED OUTPUT PLACE PAPER ON LAST LINE OF PAGE AND PRESS RETURN.
TO QUEUE OUTPUT TO DISK FILE FOR30.DAT (E.G., TO PRINT LATER
ON SPECIAL FORMS), TYPE THE WORD disk AND PRESS RETURN.

                                     ----------------------------------
PROBLEM NUMBER 2.18                  | E\J |   B̄   |   B   |  TOT  |
                                     ----------------------------------
LEGEND B                             |  Ā  |  .835 |  .165 | 1.000 |
                                     ----------------------------------
STUDENT NUMBER: 17                   |  A  |  .271 |  .729 | 1.000 |
                                     ----------------------------------
                                     |     |       |       |       |
                                     ----------------------------------


                                     ----------------------------------
PROBLEM NUMBER 2.19                  | E\J |   B̄   |   B   |       |
                                     ----------------------------------
LEGEND B                             |  Ā  |  .893 |  .380 |       |
                                     ----------------------------------
STUDENT NUMBER: 17                   |  A  |  .107 |  .620 |       |
                                     ----------------------------------
                                     | TOT | 1.000 | 1.000 |       |
                                     ----------------------------------


PROBLEM NUMBER 2.20      LEGEND B      STUDENT NUMBER: 17
 -----------------------------------------------------
| STAGES | P(B̄/Ā)  P(B/Ā)  P(B̄/A)  P(B/A) |
 -----------------------------------------------------
|    1   |  0.792    0.219    0.208    0.781  |
|    2   |  0.835    0.271    0.165    0.729  |
|    3   |  0.846    0.288    0.154    0.712  |
|    4   |  0.849    0.293    0.151    0.707  |
|    5   |  0.850    0.294    0.150    0.706  |
|    6   |  0.850    0.294    0.150    0.706  |
 -----------------------------------------------------

PROBLEM NUMBER 2.21      LEGEND B      STUDENT NUMBER: 17
 -----------------------------------------------------
| STAGES | P(Ā/B̄)  P(A/B̄)  P(Ā/B)  P(A/B) |
 -----------------------------------------------------
|    0   |  0.950    0.050    0.583    0.417  |
|    1   |  0.907    0.093    0.419    0.581  |
|    2   |  0.893    0.107    0.380    0.620  |
|    3   |  0.888    0.112    0.369    0.631  |
|    4   |  0.887    0.113    0.366    0.634  |
|    5   |  0.887    0.113    0.365    0.635  |
|    6   |  0.886    0.114    0.365    0.635  |
 -----------------------------------------------------

CPU time 1.25   Elapsed time 54.06
@
```

BEST COPY AVAILABLE

Fig.4 Instructors' Retrieval Routines

```
@EXECUTE CHAP9
LINK:   Loading
[LNKXCT CHAP9 execution]

 PROGRAM TO PRINT ANSWERS TO PROBLEMS IN CHAPTER  9

ON DISK ARE LEGENDS A THROUGH E, EACH CONSISTING OF
160 STUDENT CASES OF 32 RANDOM NUMBERS EACH

TYPE IN LEGEND LETTER (LIMITS: A TO Z) A
TYPE IN STARTING AND FINISHING STUDENT NUMBERS,
SEPARATED BY A BLANK  134 134
TYPE IN STARTING AND FINISHING PROBLEM NUMBERS,
SEPARATED BY A BLANK  1 4
FOR REPORT IN STUDENT-MAJOR ORDER, TYPE STUDENT ELSE TYPE PROBLEM
FOR PRINTED OUTPUT, TYPE THE WORD print AND PRESS RETURN;
FOR TYPED OUTPUT PLACE PAPER ON LAST LINE OF PAGE AND PRESS RETURN.
TO QUEUE OUTPUT TO DISK FILE FOR30.DAT (E.G., TO PRINT LATER
ON SPECIAL FORMS), TYPE THE WORD disk AND PRESS RETURN.


PROBLEM NUMBER 1

x    y"     f

2  2.965    2

3  2.403    4

4  3.470    5

5  5.454    5

PROBLEM NUMBER 2

SUM x =  61.  SUM y = 60.16  SUM xy =  246.44  SUM x SQ = 249.

y(x) =  -0.202 +  1.039x

PROBLEM 3          ELASTICITY=  1.054

PROBLEM # 4   STUDENT # A134

  1  -----
  2  0.375
  3  0.250
  4  0.300
  5  0.300
  6  -----
  7  -----
CPU time 1.14   Elapsed time 23.16
```

BEST COPY AVAILABLE

Fig.5 Another Answer Display

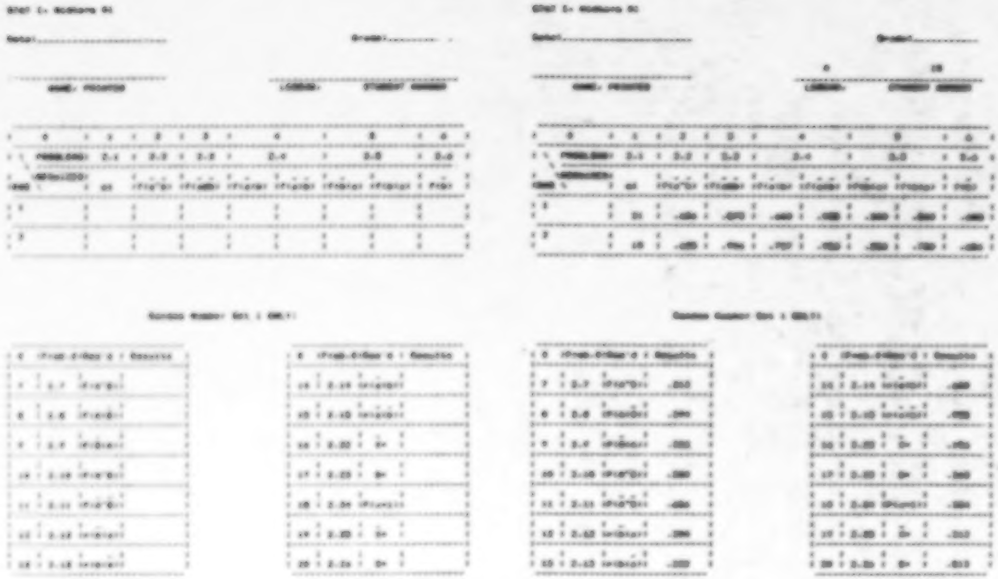


Fig.6 An Exam Blank and an Individual Answer Key (reduced in size)

BEST COPY AVAILABLE

```
  LEGEND  A     STUDENT NUMBER:     3

RANDOM NUMBER SET 1:

 145  85   5 106 110  81  50  51  79 112  87  58 147  44 138  62
 129  65  66 140 143 113  53  28 115  56  52 122   1  69  76 150

RANDOM NUMBER SET 2:

  46 131  22 149  65  50  79  58 104   6  35  21  47  94  52  42
  34 110   2  72 122 150 146  18  31 108 106  83 117  40 136 111
```

```
  LEGEND  A     STUDENT NUMBER:     4

RANDOM NUMBER SET 1:

  85  32   5  22  99  28  63 108 128 145 107  19  97   2  12  59
  26  29  84  56 100  43  90 106 149  73  69  61  82  36   1 136

RANDOM NUMBER SET 2:

  46  50  51  34  91 108 126 130  96  88  19  81  77  15  62 127
 112  20 145  63  57 125 144 133  70 142   8  71  68 131 114 101
```

BEST COPY AVAILABLE

```
  LEGEND  A     STUDENT NUMBER:     5

RANDOM NUMBER SET 1:

 113  78  15 123  59  35  38  62 103  53 147  30  36  37  74 125
  61 133  34 101 141 107   1  64  90 110 142  82  48  26   3 106

RANDOM NUMBER SET 2:

  40  29  86  14  74 117 132 139  65 134  30  35  39 114  10  70
  43  60  44  84 106  27 103  82 147 150 112 142   8  76  61  56
```

Fig. 7 Random Numbers

TOPIC 9: TRAINING STATISTICAL PRACTITIONERS

ORGANISER: PROFESSOR M BENYAKLEF

TRAINING OF STATISTICAL PRACTITIONERS

P Osinski, Munich Centre for Advanced Training in Applied Statistics for Developing Countries of the Carl Duisberg-Gesellschaft (CDG)

INTRODUCTION

The training of statistical practitioners is an important issue for both developed and developing countries. However, this paper will only deal with the training of statisticians from developing countries. It is felt that statistical offices in developed countries have the means and the possibilities to train their own staff, whenever necessary, to cope with new tasks or to introduce new methods of data collection or analysis. Newcomers to the statistical service are always trained "on the job" and most heads of offices would agree that it takes some time to transform a learned statistician into a statistical practitioner.

In most developing countries this problem is of even greater importance. In recent years a number of specialized schools and institutes have been created in LDC's where young people can obtain a degree in statistics after 2-3 years of study. Statistical theory and mathematical methods are important subjects in the curriculum of these schools and it is not at all certain that the diploma obtained will enable the young statistician to cope with the day-to-day work when joining a statistical office. Hence, there must be some additional practice-oriented training to supplement the theoretical knowledge acquired at statistical schools or universities.

While the issue of "academic" and "practical" training remains controversial the fact is that the main efforts of international organizations and some LDC Governments are directed towards "academic" training leading to a degree in statistics. The Statistical Training Programme for Africa (STPA) fostered by the Economic Commission for Africa (ECA) is mainly meant to support statistical schools in Africa. Relatively little attention is paid to the holder of a diploma of statistics once his studies are terminated. Occasional workshops on specific items of statistics organized internationally indicate the need for further training but they are completely insufficient to solve the problem.

The main concern of this paper is to outline some ideas on how statisticians after having terminated their basic statistical education can most effectively be trained (a) to cope with the tasks they are confronted with and (b) to keep abreast of new developments in statistics. The paper draws heavily on the experience the Munich Centre has gained since 1973 in organising training courses on specific subjects of statistics. Reference is also made to the author's paper "Follow-up courses", presented at the Fourth Meeting of Directors of International Statistical Training Centers - Rabat, Morocco, May 23-26, 1979.

CHARACTERISTICS OF TRAINING STATISTICAL PRACTITIONERS

The need for improving the quality of statistical personnel is a recognised and long-standing aim of practically all LDC statistical offices. This stems mainly from the request for more and better statistics made by both national and international users. In order to cope with the demand put upon them heads of offices have undertaken considerable efforts in this respect which are starting to pay off increasingly. It should, however, be noted that the training of statisticians does not always figure prominently on the priority list of LDC Governments.

The methods available for improving the skills of statisticians can be summarised in the following way: On the job training (formal or informal), training outside the office in specialized institutions or training centres and a combination of both.

Training on the job is the method widely used in developed and developing countries. It applies to both newcomers to the office and to personnel already employed. Informal training consists of a person being introduced to his job by his superior or supervisor. There is no defined training programme and the trainee learns by actually doing certain jobs which hopefully become progressively difficult. The training is strictly individual and its success depends largely on the activity and the willingness of the supervisor to transfer his knowledge. An advantage of this type of training is that the trainee learns only methods relevant to executing his job. However, there is also a certain danger that the training may become one-sided, especially in offices where the various tasks are clearly separated. This would call for a rotation programme within the office in order to avoid over-specialisation of personnel.

Formal on the job training requires a much higher degree of organisation. A programme has to be developed, the categories of people to be trained have to be defined and serious thought has to be given to the length of the courses envisaged and the availability of qualified teaching staff. Participation in training courses - although normally on a voluntary basis - should have priority over other office work to make sure that people are actually released from their duties without difficulties. The simplest form of this kind of training are courses organized from time to time on an irregular basis. A more sophisticated form used in the Federal Statistical Office of Germany consists of breaking down a rather voluminous training programme into several parts which staff members aiming at a post of responsibility are supposed to go through within a period of two years. The advantage of formal training is that a great number of young people can participate receiving all the same instruction. On the other hand course participants may feel that some of the topics of the programme are not very relevant to the work they are presently doing. This underlines the importance of a well-conceived and practice-oriented training programme.

While the conditions for executing a staff training programme exist in developed countries this is not the case in most of the LDC statistical offices. There is a lack of experience with such programmes, qualified teaching staff is in short supply and the material prerequisites such as lecture notes, background literature etc. are missing.

Under these circumstances many of the LDC statistical offices have recourse to training their staff in specialized institutions or training centers, in order to upgrade their knowledge in certain fields of applied statistics. Short-term courses are usually organized by the regional commissions of the U.N. or by institutions created in certain donor countries such as France, Germany, Netherlands, UK, United States and Japan. Most of the courses are held in the developed countries concerned, but more recently there is a tendency to organise them also in developing countries preferably on a regional basis.

What kind of possibilities does the director of an LDC statistical office have to upgrade the knowledge of his personnel in fields like agricultural or industrial statistics or national accounts? Courses in these topics are not readily available in developing countries nor are they in developed countries. What remains is to attempt a sort of "on the job training", however difficult it may be, and to use outside institutions for further training whenever the occasion arises.

In this context it should be mentioned that the upgrading of the level of statistical academic education is also widely operated by LDC statistical offices. In general the staff member is released from his duties for a period of 2-3 years with the obligation to return to the office after the termination of his studies and to work there for a certain period of time. If not doing so, staff members have to pay back the expenses incurred by the Government for their studies. The observance of this rule, however, is difficult to enforce.

PURPOSE OF TRAINING

A statistician participating in a course of short duration on a specific type of statistics is likely to be interested primarily in items having direct relevance to his job. He rightly expects that his efforts will enable him to have a better insight into the problems he is confronted with in order to do a better job. The first and overriding consideration when organising such courses, therefore, is that they are practice-oriented.

This is easier said than done. Most people will admit that the borderline between "theory" and "practice" is artificial and that a certain amount of theory has its place in any course programme, however practical it may be. On the other hand, a practical course can never be expected to cover all facets of reality. Many statisticians participating in a course expect just this to happen. The ideal for them would be to go home with a sort of recipe book listing all the steps to be undertaken in a statisical survey and giving detailed advice on how to deal with all possible problems. The trouble is that local situations in LDC's are extremely variable and do not lend themselves to ready-made solutions.

The aim of training should, therefore, be to upgrade the specialized subject-matter knowledge of the statistician enabling him to take decisions in an independent and well conceived manner whenever problems arise. To do this he needs a thorough knowledge of the concepts and definitions used in his special field of statistics and some background information why certain concepts or methods have been given preference as compared to other possible solutions. The exchange of experience among statisticians is also very important in order to learn from each other and to adopt methods of data collection, treatment, publication and analysis which already have been used successfully.

SCOPE OF TRAINING

When organizing training courses it is of vital importance that the topic selected is well defined. In principal one can choose either a methodological approach dealing with selected statistical methods applicable to all fields of applied statistics or a sectoral approach concentrating on specific subject-matter such as agricultural statistics, industrial statistics etc. The latter method has the advantage that it corresponds to the organizational set-up of most statistical offices in LDC's. Typical items to be treated are data collection methods (censuses, sample surveys), questionnaire design, data processing with special reference to EDP, data analysis and publication. While the topics are treated in more general terms when using the methodological approach, teaching is geared to the specific subject when using the sectoral approach. The interrelationship of the subject-matter statistics under consideration with other statistics should also be shown in order to rouse the awareness of the statistician that the statistics he is doing are an integrated part of an overall framework and hence need coordination with related statistics.

When dealing with statisticians from LDC's it should be kept in mind that many statistical offices are still badly equipped with both manpower and material. Therefore, it makes no sense to confront statisticians from these countries with the most sophisticated methods of data analysis or publication as there is no chance that they will be able to apply them in the near future. This does not mean that very new and advanced methods should be completely excluded from the lectures, but main emphasis should undoubtedly be on methods which can be readily applied.

Another question is whether related subjects should also be included in a training programme. When dealing with economic statistics some selected lectures on the functioning of the economy would certainly be useful. The study of population statistics may also be supplemented, for instance by lectures on sociological aspects of vital statistics. It is somewhat difficult to decide how far one should go in including related subjects in a statistics course. It is clear, however, that a profound knowledge of the subject area a statistician is studying (agriculture, industry, prices etc.) will permit him to produce better statistics and to analyse the results more thoroughly.

An interesting attempt to statistically analyse living conditions and the economic decision making process in African rural societies is undertaken by a group called AMIRA created by the "service de cooperation" of the French statistical service INSEE. This inter-disciplinary group includes inter alia statisticians, economists, sociologists, anthropologists. A similar approach is followed by the Development Research Center of the World Bank and its study of living standards measurement. This integrated approach to statistics has its merits with regard to well-defined research projects at a village or rural level and may lead to a better understanding of household statistics at a regional level. The Munich Centre's experience with presenting some of the findings of the group to our African students has been somewhat disappointing in so far as most of them regarded the situations described as "old-fashioned" and due to disappear in the course of modernisation of the society.

VARIOUS APPROACHES TO TRAINING

The purpose and the scope of training being fixed it has to be decided which method to use in order to motivate the statisticians as much as possible. Various approaches have been outlined by Mr. Waast in a meeting called "journées pédagogiques" organized in Paris by the "Centre Europeen de Formation des Statisticiens Economistes" (CESD) from the 7th to the 11th of September 1981. About 60 persons from statistical schools, training centres and organizations fostering the training of statistics participated. Mr. Waast especially referred to four approaches:

- courses on special statistical topics
- courses on statistical methodology
- critical evaluation of surveys
- practical execution of a survey

Courses on special statistical topics are still the most commonly used method of training statisticians. Since 1973 the Munich Centre has organised two courses of 4 months' duration annually, one of which is in English and one in French. The subjects dealt with so far are national accounts, foreign trade statistics, industry and handicraft statistics, demographic statistics, agricultural statistics, household statistics, social statistics. Around 500 statisticians from African and Asian countries having a university degree or a similar training in statistics have participated in the courses.

In their most primitive form courses on special statistical topics have mainly two disadvantages: they concentrate exclusively on the subject-matter under discussion without showing the interrelationship with other statistics and they are in the form of "ex cathedra" lectures leaving little room for discussion and exercises. Attention should, therefore, be given to exercises being included in the course programme and showing the use to be made of the data collected and their place within a coordinated framework of statistics. The lecturer should be instructed to permit questioning at any moment in order to initiate a dialogue with the statisticians attending. This will give them the opportunity to discuss the actual problems of their daily work, to exchange their experiences and to ask for potential solutions. As far as possible the lecturer should aim at using a fair amount of the time available for practical exercises in order to check that the points made were correctly understood. If these items are observed, courses on special statistical topics are a useful tool for training statisticians.

It should, however, be mentioned that much depends on the lecturers available. Experience in statistical work in LDC's is of vital importance in order to give advice on how to adapt conventional concepts to local conditions. The preparation of lecture notes and written exercises beforehand is also of great value.

Courses on statistical methodology have quite a different starting point. A topic of general economic interest such as an agricultural development plan, or the fight against inflation, may be chosen. First the problem has to be analyzed, then the need for data relevant to the subject has to be determined, and finally the best way of obtaining the data needed has to be elaborated. The course may include a discussion on the validity of existing statistical instruments (national accounts, multi-purpose household surveys etc.), data collection techniques (sample surveys, interviews etc.), their application in a given social environment and the need for developing an information system.

The approach is a rather theoretical one. It is interdisciplinary in the sense that it may embrace all kinds of statistics. Formal lectures are kept to a minimum and emphasis is placed on the comments and proposals made by the participants. There are a number of disadvantages. One, certainly, is that the group should be restricted to keep the discussion within reasonable limits. More important still is the fact that the participants need a rather advanced technical knowledge of the problems under discussion in order to be able to cooperate fully. This approach, therefore, is more suited to rather short seminars for already experienced statisticians.

A third way of training statisticians consists of the critical evaluation of surveys already undertaken. The exercise includes a study of the preparatory documents (objectives of the survey, methodology) and the survey documents (questionnaires, instructions to interviewers, classifications etc.). The method has the advantage of being based on documents used for real work. In order to stimulate the interest of the participants it would be desirable to investigate a series of surveys using different methodologies rather than a single one. It may, however, be difficult to bring together the necessary background material.

The study of documents may be beneficial to students, especially the questionnaires used and the way they are designed. The analysis of other documents - mostly rather lengthy ones - is often not very stimulating and it is difficult to capture the attention of the course participants by going through them. The course tends, therefore, to turn into a series of lectures on the documents presented by the responsible lecturer rather than an analysis made by the students. A distribution of the documents to the participants may, however, be of great value to further reference.

The last approach mentioned concerns the practical execution of a survey. There is no doubt that a small group of statisticians working together in a team to deal with all aspects of a survey is a good way of making them acquainted with the difficulties of applying statistical methods to real situations. The method educates to be imaginative, to compromise and to find solutions to unforeseeable problems. On the other hand it is time-consuming and rather costly. In order to be of some general value the number of cases included in the survey should be sufficiently large and the sampling frame correctly conceived.

While the method has certain advantages it should be noted that its application very often encounters difficulties. A certain infrastructure is indispensable and the social environment must lend itself to surveys. The training of LDC statisticians in European training institutes by applying this method is generally hampered by restrictive regulations concerning surveys, a certain unwillingness to respond to inquiries on a voluntary basis and a growing concern about the confidentiality of individual data. However, the organisation of surveys representing undoubtedly a very important aspect of work in a course on applied statistics does not dispense with studying more "theoretical" items such as concepts and definitions or analytical methods.

A useful tool for training statisticians is case studies. The US Bureau of the Census, through its International Statistical Programme Center (ISPC) has developed a set of materials, i.e. "Atlantida: A Case Study in Household Sample Surveys". It is a rather complex study dealing with all phases of a survey programme and emphasizing the practical aspects rather than the theoretical ones of household surveys. More recently the French statistical office INSEE has developed some case studies. The mythical country "Baramu" represents a case study centered on employment, education and health problems. The intention is to enlarge the study to comprise also agricultural statistics, foreign trade and industrial statistics. Another case study "Kangare" shows how to build up national accounts from given data. The case study "Bernard" intends to provide an introduction to business accounting. The French "Centre de Perfectionnement Economique et Financier" (CEPEF) mainly uses two case studies namely "La Camesie" dealing with annual forecasts and "Tercera" centering on the planification of rural development.(*) All case studies try to approach reality as far as possible. Their educational value depends very much on the ability of the teaching staff to present the studies in an interesting way, to guide the discussions of working groups and to give advice on the various problems as the students go along. The teachers should, therefore, be very qualified having also a rather profound experience in LDC's.

A reproach often made to courses on special statistical topics like, for instance, population statistics is that these courses tend to compartmentalise the teaching of statistics. However, the danger of doing this appears to be even greater when evaluating or executing a survey. The latter exercises do not necessarily require a discussion on the use to be made of the data and the interrelationship with other statistics. In all methods used there is the danger that, for instance case studies or evaluations turn into "ex cathedra" lectures rather than being the result of efforts undertaken by the students.

(*) For more detailed information cf. Compte-rendu des journées pédagogiques des écoles de statistique africaines, Paris, 7-11 September 1981. "Centre Européen de Formation des statisticiens économistes des pays en voie de développement", 3 Av. Pierre Larousse, 92240 Malakoff, France.

Whatever the method of training, one should always keep in mind that a statistician participating in a training course is mainly interested in topics which are useful to the work he has to do in his office. The concentration on a well-defined subject, therefore, still appears the best way to stimulate his cooperation. He will also be receptive to learning which use can be made of the statistics he produces and the interrelationship with other statistics. A combination of lectures, discussions and case studies appears to be the best way of conducting a "practical" training course, whereby a lot of room should be devoted to problems brought forward by the students themselves. Whenever the teaching can be combined with the execution of a real survey this opportunity should be used, as the experience in the field is of greater value than any theoretical case study.

MATERIAL AIDS TO TRAINING

The conventional blackboard and the overhed projector are still the most commonly used aids to training. The direct contact with the lecturer, the possibilities of asking questions and of opening a dialogue on unclear points are valuable assets for the participants. On the other hand the response of the audience enables the lecturer to adapt the progress of his lecture to the understanding of the students. It may be that these advantages had a somewhat retardatory influence on the use of video tapes. It may, however, be expected that they will play a more important role in the future due to technical progress making their use relatively cheaper. The availability of suitable programmes is another important precondition.

The preparation of lecture notes and the distribution of an annotated agenda before the lecture is also of great importance to the success of training. The course participants should know at any moment which point of the lecture is discussed and how it relates to what has been said beforehand and to what is still going to follow. Lecture notes should summarise the content of the lecture in concise form. They are very important as reference material for further study and as aid-memoire for the participants, especially if they come from LDC's where relevant literature or experienced discussion partners may not be available.

Microcomputers have recently become very powerful and extremely cheap, so that even smaller and financially poor statistical offices can afford them. The use of microcomputers should, therefore, become a standing feature in any course programme for statisticians. The purpose should not be to convert them into programmers or computer experts, but rather to familiarise them with the use of electronic equipment. This will hopefully contribute to a better understanding between statisticians and EDP technicians and facilitate their cooperation. In this sense it is not important that the statistician knows a series of commonly used computer languages. Basic training should be centered on supplementing the statistical courses by practical exercises on the computer. The course participant should be given the opportunity to manipulate the microcomputer as often as possible and to learn its use by the "trial and error" method. This will enable him to open an informed dialogue with the EDP technician in case the job on hand requires the use of a big computer.

In this context some general remarks on the use of computers may be in order. At the moment each office or institution of some reputation will try to have its own computer - and the bigger the better. The accounting work has then, of course, to be adapted to the requirements of the computer which in many cases leads to substantial organizational changes and sometimes to rather significant additional paperwork. The progress in technology increasing the capacity of microcomputers enormously now makes it possible to use these devices for jobs of rather substantial dimension. For the time being producers are still more concerned with producing sophisticated hardware leaving the software aspects somewhat aside. But things are changing quickly and more and more software for microcomputers becomes available. Therefore, it is quite legitimate to expect that a series of microcomputers, each one well adapted to specific jobs, may be able to replace a big computer in the foreseeable future. This is especially important to LDC's statistical offices where, for a number of reasons, the use of big computers encounters difficulties.

TEACHING STAFF

The teaching staff for training statistical practitoners could consist of permanent staff members of the institute doing the training, part-time lecturers or a combination of both. All depends on whether the training is continuous or sporadic and whether the training is supply-oriented (repetition of a fixed set of subjects) or demand-oriented (changing subjects on request).

Supply-oriented training institutes organizing courses on a regular basis tend to have permanent teaching staff. As the subjects dealt with are more or less the same in each period each staff member has his responsibilities within an overall training programme. The advantage is that coordination among the staff members is relatively easy. On the other hand the system is costly and budgetary difficulties are likely to arise in connection with requests for new staff. It is therefore more suited to education in statistics rather than to training practitioners.

If the training is demand-oriented, i.e. the subjects of courses are constantly changing it is preferable to employ part-time lecturers. Best suited are experienced statisticians from statistical offices or institutes as well as academic people with experience in practical statistical work. This will ensure that the training has a certain bias towards practical issues which is rather desirable when training practitioners. Experienced statisticians serving as part-time lecturers usually have a profound knowledge of the subject-matter, they know the latest developments in their field and are able to advise on practical issues. The use of part-time lecturers ensures a certain flexibility and is financially less burdensome. On the other hand it is not always easy to put together the "right" team for a course as good statisticians tend to be rather busy. Coordination is also difficult and lecture notes are often a problem.

When training statisticians from LDC's a knowledge of local conditions is indispensable. Commonly used concepts in developed countries are not always transferable to LDC's and advice on problems brought forward by the course participants will hardly be possible without LDC experience.

COST OF TRAINING

The cost of training is an important consideration in any training programme. In general the opinion prevails that training courses organised in LDC's are cheaper than those held in developed countries. This may be true to a certain extent, but expectations in this respect are often higher than real savings. All depends on whether the organiser of a course is able to ensure the co-operation of a local school, institute or university which is prepared to provide the infrastructure needed (classrooms, teaching aids, accommodation, sometimes even part-time lecturers). If this is the case courses in LDC's can indeed be organised at relatively low costs. However, if course participants have to be lodged in hotels, classrooms etc. hired at normal rates, UN per diem rates paid to participants and teachers flown in from abroad, the saving made on lower expenses for airplane tickets for the students may be largely offset.

Most statistical training institutes are subsidized in one way or the other. The expenditures for housing, administrative and teaching staff, materials etc. are often paid from Government aid funds. Costs vary widely depending on the organization of the centres and the permanent staff employed. It is therefore difficult to establish any general rules. However, it can be said that training costs are enormously high if the organiser intends to recuperate all costs by tuition fees.

At present the direct costs for a statistician participating in a 4 months training course at the Munich Centre are approximately as follows:

-	air economy ticket	US $ 2,000.--
-	clothing allowance	US $ 250.--
-	books	US $ 100.--
-	fellowships ($ 500 per month)	US $ 2,000.--
		US $ 4,350.--

EXPERIENCE OF THE MUNICH CENTRE

Since its inception in 1973 the Munich Centre for Advanced Training in Applied Statistics for Developing Countries has organised the following courses:

- 2 four months courses annually in Munich one of which is in English and one in French
- Seminars of 3 weeks duration in Munich for directors of statistical offices or their deputies
- 4 to 6 weeks courses in developing countries

The 4 months courses are meant for statisticians from LDC's having an academic degree in statistics or an equivalent education and some years of experience in practical statistical work. In spite of efforts to keep the group as homogeneous as possible we are not always successful even though attention is paid to arriving at the same educational background. The professional experience of a statistician is decisive for his performance and much depends on the office he is working in. The maximum number of course participants is 23 coming from about 18 different African or Asian countries. Hence, a certain variety of qualification is unavoidable.

More recently a certain change has been noticeable in the composition of the participants. In the early 1970's most participants had received their basic statistical education from European universities. Nowadays practically all participants studied in their home countries. While this may not necessarily affect their theoretical statistical knowledge, a certain decline in language proficiency is to be noted. Hence, it is advisable to request a language test before admission.

Up to now the course programme was always centered on a specific topic of statistics such as industrial statistics, agricultural statistics etc. The concentration on one topic does not mean that it is looked at in isolation. The course duration of four months permits going through the whole process of statistical surveys, from data collection through publication and analysis. Special attention is paid to the use of data, their interrelationship with other statistics and their usefulness for national accounting and planning purposes.

In 1983 a new approach will be tried by dealing with the topic "collection and analysis of data - methodological aspects". It is mainly meant for statisticians having the function of preparing surveys or analysing data from a variety of sources.

The Munich Centre works with part-time lecturers. They are experienced statisticians who are able to teach in English or French thus avoiding interpretation. Most of them have extensive experience in LDC's. Course participants from LDC's have a preference for lecturers from their own countries. Whenever possible the Munich Centre invites LDC lecturers in spite of the sometimes burdensome additional travel cost. The main problem with part-time lecturers lies with the coordination of their interventions and the preparation of lecture notes which sometimes never arrive due to other obligations of the lecturers.

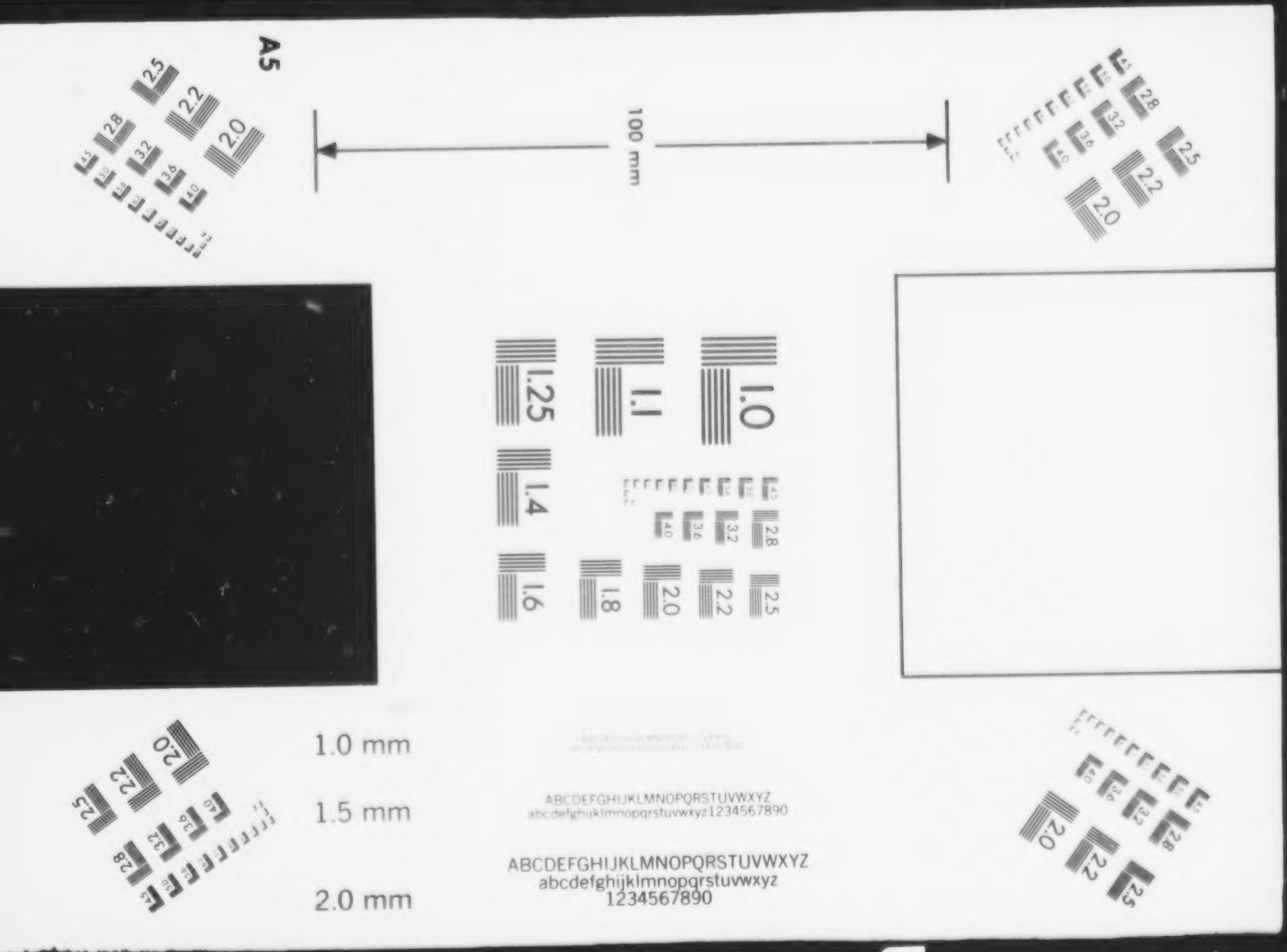
The teaching is informal. Lecturers can be interrupted at any moment for comments or questions. The cooperation of the course participants in communicating their experience is essential to the success of a course and should be encouraged. Practical exercises are highly appreciated by the participants and should form a substantial part of the programme. Lectures in the morning and discussions and practical work in the afternoon may be a workable alternative approach.

Seminars for the directors of statistical offices or their deputies were organised on an irregular basis. The topics were "the organization of a statistical service" and "the setting and implementing of statistical priorities". The duration was three weeks. This is definitely too long and should be shortened to two weeks as it is difficult for the directors to be absent from their office for longer periods. In preparing seminars of this kind the difficulty lies with finding suitable topics of interest to the directors without being too technical. The various items on the agenda should be well prepared by background papers distributed beforehand. Care should be taken to ensure the cooperation of internationally well-known statisticians to prepare the papers and to serve as discussion leaders as this will have a positive impact on participation. The greater part of the seminar should consist of discussions and exchange of experience. A report reflecting the findings should be prepared by the participants.

Courses in developing countries have been carried out by the Munich Centre since 1980. The topics were "Household statistics" (Cameroon), "EDP for statisticians" (Thailand), "Agricultural statistics" (Ivory Coast), and "1983 World Census of industrial statistics" (Ethiopia). In each case the cooperation of an already existing institute or organization could be ensured, i.e. the

"Institut Statistique de la planification et de l'economie appliquee (ISPEA)" in Yaounde, the "Asian Institute of Technology" in Bangkok, the "Ecole statistique d'Abidjan" and the ECA in Addis Ababa. The provision of class-rooms, secretarial help, typing and reproduction facilities, accommodation therefore did not represent any special problems. In all cases it was also possible to employ local part-time lecturers. This ideal solution works, however, only during the summer vacation when the regular students of the schools are absent. At other periods the schools work at full capacity and cannot accommodate additional training courses.

LDC's unconditionally support training courses organised in their countries. Administrative procedures concerning nominations appear to be easier and wider publicity is normally given to the courses. It is advisable that the organiser limits participation to a regional level. This ensures a certain homogeneity of the participants and diminishes travel cost. Another important advantage is that practical work can sometimes be organised by conducting a real survey.



1.0 mm

1.5 mm

ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz1234567890

2.0 mm

ABCDEFGHIJKLMNOPQRSTUVWXYZ
abcdefghijklmnopqrstuvwxyz
1234567890

PREPARING STATISTICIANS FOR GOVERNMENT SERVICE -
A VIEW FROM THE UNITED STATES

Marie D Eldridge, United States Department of Education
and Katherine K Wallman, Council of Professional Associations on Federal Statistics

INTRODUCTION

In 1979 the American Statistical Association (ASA) Section on Statistical Education formed a Committee on Training of Statisticians for Government. The establishment of this Committee was an outgrowth of work by a predecessor group (Committee on Training of Industrial Statisticians 1980) sponsored by the ASA. Under the leadership of the authors of this paper, the Committee on Training of Statisticians for Government completed its assignment and presented its findings on education and training for federal statisticians at the 1981 annual meeting of the American Statistical Association. The full report (Committee on Training of Statisticians for Government 1982) was designed to assist colleges and universities in developing their statistics curricula, to serve as a counselling tool for prospective federal government statisticians as they select their training courses, to provide guidance to managers who recruit and supervise statisticians as they select employees and plan for their career development, and to offer recommendations on qualification standards for federal government statisticians. Material in the present paper is drawn largely from the Report of the Committee on Training of Statisticians for Government; the paper focuses on preservice and inservice education and training for government statisticians, and on recommendations to colleges and universities and the national government.

THE FEDERAL STATISTICIAN: A PROFILE

Perhaps the largest single employer of statisticians in the United States is the federal government. As of October 31, 1978, the federal government employed over 3,900 statisticians, as well as 3,240 statistical assistants and numerous other employees performing statistical duties but classified in different job series (U.S. Civil Service Commission). While the federal government has an extensive need for statisticians with highly specialized capabilities, the majority of the government's statisticians are generalists involved with data collection activities. Even the Bureau of the Census, with its major emphasis on demographic analysis and research in statistical methodology and statistical data processing techniques, has the collection of general-purpose statistical data as its primary purpose.

* Contribution of the U.S. Government, not subject to copyright in the United States.

The Survey Statistician

Although survey statisticians as a group comprise the major segment of federally employed statisticians, the responsibilities of individual employees differ substantially. A junior or midlevel statistician working on a complex survey for an agency such as the Bureau of the Census may perform duties solely related to frame maintenance, sample selection, survey estimation, data adjustment, quality control, or data analysis. In the Department of Defense, by contrast, the entire survey process from design to analysis is likely to be done by one or two individuals.

The size of the agency responsible for a survey is also a key factor in determining how and by whom the survey is designed and the data are collected. In an agency with few statisticians, the survey design is frequently done by an outside contractor and the actual data collection is virtually always done on contract. In these cases, the survey statistician's responsibilities are largely concerned with writing requests for proposals, evaluating contract proposals, and monitoring the performance of contractors. In most of the large general-purpose statistical agencies (Bureau of the Census, Bureau of Labor Statistics, the Department of Agriculture's Statistical Reporting Service, and the National Center for Health Statistics), the design is virtually always done by an internal unit; the data collection is then done by field units (enumerators) across the country.

The nature and extent of the federal statistician's duties are affected by his or her professional level, as well as by the size of the employing agency and the complexity of the survey. At the junior professional level, much of the statistician's time is devoted to on-the-job training in all aspects of his or her office's program. In addition, the junior statistician frequently becomes an expert in some specialty - either in some analytical technique (such as data adjustment), in some procedure (such as contracts or publications), or in some topic (such as an industry).

At the middle professional level, the survey statistician is expected to be able to handle non-routine technical problems. Responsibilities include developing procedures and schedules for data collection and writing operating manuals and procedures. Technical expertise is used to help define concepts or to identify problems and to develop possible solutions. In the smaller agencies, the midlevel statistician may have full responsibility for the entire survey process from planning the design to writing the final report.

The senior level survey statistician's time is generally divided among supervision, administration, and program planning. The beginning supervisor generally is responsible for a unit comprising three to ten people (statisticians, subject matter specialists, statistical assistants, and secretaries) and administers one or more surveys or programs. In addition to planning the unit's activities, the senior statistician participates in planning the office's overall program and in discussing budget requirements with budget officials. Representing the agency in dealing with personnel at the highest levels in other agencies and in private firms and the training of agency personnel are usually the senior statistician's responsibility.

Like the senior statistician in charge of data collection or validation and the production of estimates, the senior statistical analyst spends a great deal of time in program planning, supervision and administration, but unlike the senior survey statistician he or she may devote a large portion of time to applied research. This research will be either program-oriented research - that is, analysis of the output from a particular statistical program - or topic-oriented research, using a wide variety of data sources and a wider variety of approaches and techniques to analyze a particular topic, such as a comparison of cost-of-living indices, techniques for analysis of quality change, or a study of the economic benefits of education.

Demographers, Biostatisticians and Other Specialists

While the majority of U.S. federal statisticians are involved in general data collection activities, the government also employs many statisticians whose jobs are highly focused in other areas. Such specialists include demographers, biostatisticians, engineering statisticians (i.e., specialists in such areas as quality control and reliability), researchers in statistical methodology, and an increasing number of specialists in computer software for statistical applications.

Demographers in the federal service work on various programs and problems including public health, manpower analysis, family planning programs, immigration and emigration policy, population forecasting, estimation and projection, and other areas. The principal federal employer of biostatisticians is the National Institutes of Health, which is the medical research arm of the Public Health Service. Biostatisticians work in lesser numbers for other Public Health Service agencies, serving principally as epidemiologists. In the National Center for Health Services Research, biostatisticians focus on analyses of health services data, in the Food and Drug Administration on the evaluation of drug safety, and in the National Center for Health Statistics on the analysis of vital and health statistics.

Statisticians with specialized education in assurance sciences and experimental design are also in demand in the government. Quality control, reliability, availability, and maintainability are all aspects of assurance science used by Department of Defense statisticians in the course of designing, procuring, storing, and using weapons systems and related material.

Scattered throughout the government are statisticians whose primary function is statistical consulting with scientific researchers. Areas of application include the physical and engineering sciences, agricultural sciences, environmental and geophysical sciences, and the biological and medical sciences. These statisticians work in close collaboration with scientific specialists to develop empirical models of phenomena and statistical models for measurement or observation processes.

The federal government also employs a sizeable number of mathematical statisticians in the area of methodological research. Typical areas of research include sample design and estimation, the theory and methods for survey evaluation, data adjustment, the improvement of data quality and survey operation through efficient quality control procedures, the assurance of respondent confidentiality in published data through mathematical disclosure analysis, the computerization of maps, computer editing, the establishment of production standards for survey personnel, and statistical modelling.

A final example of the specialized uses of statisticians in the federal government involves computer programmers. In several agencies, individuals responsible for programming complex sample designs and the associated estimates and variance calculations are classified as statisticians. Such individuals are also regularly called upon to apply a myriad of standard statistical software packages to survey or other statistical data and to develop software for special purposes.

While the job requirements of the specialized mathematical statistician are frequently much more technical in nature than those of the survey statistician described earlier, the level of responsibility of the junior and midlevel mathematical statistician is similar to that of a survey statistician employed at a corresponding grade. The primary difference in the work of the two types of statisticians occurs at the senior level, where the mathematical statistician is still commonly found in a nonsupervisory advisor position with highly specific technical demands and requirements. Almost all survey statisticians at this level, on the other hand, are in management or supervisory positions.

EDUCATIONAL NEEDS OF FEDERAL STATISTICIANS

Because of the heavy emphasis on data collection in the U.S. federal statistical system, it is virtually imperative that the effective federal statistician be well trained in survey techniques. (Exceptions to this rule are those highly specialized statisticians providing statistical assistance in a scientific research community). Beyond the earliest entry levels - where an individual would likely not need statistical training beyond the fundamentals of statistical theory, inferential statistics, and correlation/regression techniques - knowledge of survey methods is vital to most government statisticians throughout their careers. In most agencies, the lack of a solid background in applied statistical methods and sample survey procedures would hamper the effectiveness of the statistician by the time the mid-professional level is reached. To be fully prepared for the job at this level, the statistician should have had training in the areas of advanced statistical theory, applied sample survey methodology, computer programming, and technical writing.

The advanced statistical theory desired at this level would include mathematical statistics, theory of probability, matrix theory, and sampling theory. The first three courses are standard at most institutions of higher education, usually forming part of the required curriculum for a degree in statistics. Unfortuantely, sampling theory - probably the most important of the four for many government statisticians - is still not offered in some institutions with statistics curricula and is not required in many others. In fact, in some institutions the student receives almost no applied training until after the master's level. In addition, the quality and appropriateness of the sampling theory courses vary widely from institution to institution. Frequently, the courses are taught by persons who themselves have had no experience in designing a survey or specifying a sampling plan. Such courses - which are usually limited to the derivation of endless mathematical formulas for variances of alternative designs - are of marginal value to government statisticians. A considerably more useful sampling theory course would be one that is presented in the context of how one makes decisions among alternative designs, considering both cost and precision. Unfortunately, few U.S. colleges and universities offer sampling theory courses of this type. The problem is so severe that some agencies find it necessary to encourage newly hired statisticians to take an adequate sampling course.

Of particular importance to the junior professional survey statistician is education in applied sample survey methods. Unfortunately, adequate training opportunities in this area are even rarer than are adequate sampling theory courses. Most graduate statistical curricula are so theoretically based that the elementary principles of applied survey methods are not covered at all. Other than in many social science statistics courses, the principles of questionnaire design and the analysis and interpretation of survey data are rarely touched. And with the exception of a possible digression in a sampling theory course, students very rarely delve into the vital area of the estimation and control of nonsampling errors in survey data. Missing also is adequate training in imputation and other incomplete data techniques. Yet these areas represent the everyday working tools of the federal survey statistician at the junior and middle professional levels.

The final areas of technical education needed by the junior professional statistician are computer programming (particularly FORTRAN and the more common statistical software packages such as SPSS, SAS and BMD) and technical writing. Virtually all statistical curricula now include heavy reliance on the computer as a statistical tool, but technical writing is usually forgotten. As Barbara Bailar stated in her paper describing government training needs:

> The writing of statistical reports seems to be a problem because of poor writing ability. Graduates may never [have] had to write reports, especially if [their] training was primarily in mathematics. Frequently one is told that one reason graduates majored in mathematics was to avoid writing, which they say they dislike. Unfortunately, it is of little use to be able to analyze data beautifully but not be able to communicate the results of that analysis.

The need for writing skills cannot be overemphasized. As statisticians move into managerial positions, communication - both written and oral - becomes a vital ingredient of a successful performance. As a manager, the statistician must frequently communicate important, but highly complex, technical details to senior officials with no technical training. Inability to communicate effectively in such circumstances could lead to important decisions being made on an incorrect or incomplete understanding of available information.

A possible solution to this problem would be for colleges and universities to require their mathematics and statistics majors to take an applied course in which historical data are analyzed (or applied to a mathematical model, etc.) and a technical report of the study is required. This would give the student at least a nominal exposure to technical writing and applied techniques to counter in part the current imbalance of theory over practice.

By the time a federal survey statistician rises to the top of the mid-professional level, he or she is likely to be responsible for an entire survey or a segment of a major survey. It is at this time that consulting and analytical skills become essential as well. Statisticians at this level should obtain additional technical education in the areas of statistical analysis methods such as nonparametric methods, stochastic processes, time series analysis, and multivariate analysis techniques. In addition, in-service training in introductory principles of supervision is essential, as statisticians at this level are frequently team leaders, if not formally supervisors.

Finally, senior level general or survey statisticians must have a broad background in supervisory education. While their earlier technical education will still be extremely important to them, the job demands at this level turn increasingly to management.

In 1979, the Statistical Policy Coordination Committee (SPCC), established to advise and assist the President in the improvement, development, and coordination of federal statistical services, appointed a Task Force on Civil Service Reform to make recommendations on policies and activities which would contribute to ensuring the integrity and improving the performance of the federal statistical system. One of the areas considered by the Task Force was the opportunity which executive development programs offer for strengthening the skills and effectiveness of senior statistical personnel in the federal government. The 1981 Report of the SPCC Task Force on Civil Service Reform concluded, among other things, that:

1. Senior statisticians tend to develop their experience in the technical aspects of their positions, and often have been relatively isolated from broader policy and management activities such as interface with the legislative and budget processes, and
2. In many instances, senior statistical personnel in the federal service have spent most, if not all, of their professional careers within a single agency or narrow field of specialization.

As a consequence, the Task Force recommended that executive development programs for senior statisticians in the federal service focus on the acquisition of managerial skills and provide an opportunity for participants to gain a broader understanding of the policies, processes, and programs of the federal government. Types of development needed to improve the government statisticians' effectiveness as senior managers included, for example, education in communication skills, personnel management, and coordination with a wide variety of outside groups and agencies.

Acquisition of the managerial skills needed by the senior government statisticians can be undertaken through various means ranging from formal classroom instruction to informal briefings or discussion groups. Among the options suggested by the Task Force were formal academic training, assignments to another part of the agency (such as the budget or legislative division) or to a different agency for a period of two to three months, intensive (two to three week) seminars, short (one to five day) workshops, information visits or briefings, attendance at management or other policy formulation meetings, and independent study or discussion groups. The specific mix of executive development activities the statistician should pursue will depend, of course, on the individual's background and interest in assuming managerial responsibilities in the government statistical system.

A summary of the desired education for government statisticians is contained in the following table. (For additional views on this topic, see papers published by Fellegi, Malinvaud and Murthy 1978). In some specialised areas, additional courses in a particular subject may be required, or certain statistical analysis courses may be needed earlier in the employee's career.

IN-SERVICE EDUCATION AND TRAINING OPPORTUNITIES FOR FEDERAL STATISTICIANS

To meet the specialized requirements for education and training of federal statisticians, and to fill existing gaps in their educational experience, the U.S. statistical agencies have relied heavily on in-service training. Apart from the natural "on-the-job" training that derives from association with other statisticians, most federal statistical agencies make a wide variety of education and training opportunities available to their employees.

The most frequent method of in-service training used is agency support of the statistical training offered by local universities. Although many universities now offer degree programs at night to attract full-time workers - particularly in areas with large concentrations of federal workers - there are still few direct arrangements between those institutions and the federal statistical agencies. Many universities educate their doctoral students only for academic careers. Fortunately, some recent activities seem to be changing this picture.

One such program, called the Statistical Survey Institute, was established jointly by the American Statistical Association and the Statistical Reporting Service (SRS) of the Department of Agriculture in 1976. The objective of this program is to enhance graduate education and research in statistical theory by tying instruction in these areas to data collection and analysis at the SRS. A second program is a joint program of the American Statistical Association and the Bureau of the Census, partly funded by the National Science Foundation. The objective is to develop a program of research applied to the social sciences and to supplement the experience of senior social scientists and the education of graduate students in statistics, economics, demography, and related areas by exposing them to the real statistical problems encountered in a large data collection agency.

Suggested Technical Education and Training for Federal Statisticians

Level	Courses*
Entry level	Theory of Statistics
	Statistical Inference
	Correlation and Regression Methods
	Elementary Scaling Principles
	Graphical Presentation of Data
Junior Professional	Sampling Theory
	Questionnaire Design Principles
	Mathematical Statistics
	Theory of Probability
	Analysis and Interpretation of Survey Data
	Estimation and Control of Nonsampling Errors in Survey Data
	Imputation/Incomplete Data Techniques
	Matrix Algebra
	Fortran/Cobol, SAS/SPSS/BMD, Data Base Management Language
	Technical Writing
	Basic Government Contract Procedures
Mid-level Professional	Consulting Principles and Practice
	Nonparametric Methods
	Stochastic Processes
	Time Series Analysis
	Multivariate Analysis Techniques
	Advanced Scaling Principles
Senior Professional/Manager	Introduction to Supervision
	Personnel Management, Labour Relations, Equal Employment Opportunity and Related Administrative Areas
	Effective Communications for Managers
	Government Budget, Legislative, Regulatory and Procurement Procedures
	Relationships with Other Governmental Levels, Public Interest Groups, the Media, and the Private Sector

---

* List is not intended to be in priority or educational sequence order.

Some agencies have developed long-term education agreements with university departments. Under these programs, selected employees are enrolled in the regular graduate program of the institution. In most cases a master's degree is earned by the trainee, although the actual receipt of a degree is secondary (from the federal government's point of view) to the value of the training received.

A major segment of the in-service training that is available to federal statisticians is provided by the agencies themselves. Perhaps the best known of such programs is that of the U.S. Department of Agriculture (USDA) Graduate School. Two of the courses offered through this program - Advanced Agricultural Statistics and Sample Survey Methods - were specifically designed and developed to fill training voids encountered by agricultural statisticians at the entry level.

RECOMMENDATIONS

The preceding sections of this paper describe the environment in which the U.S. government statistician works, the education and training needed in order to work effectively in that environment, and in-service training opportunities for federal statisticians. In the federal government, the well-educated statistician will be one who has been exposed to a balance of theory and practical applications. The statistician who can bring appropriate statistical theory to bear on various methodological and practical problems, and translate this knowledge into terms understandable by the program manager and the policy-maker, will have the greatest impact in terms of bringing statistical expertise to bear on government programs and policies. While there are isolated needs for pure theorists, statistical applications remain the primary tool of the government statistician. For this reason, tools such as computer programming, applied sampling theory, data systems design, incomplete data procedures, technical writing and technical consulting are of primary importance.

In view of its findings with respect to the nature of government statistical work in the United States, the Committee on Training of Statisticians for Government recommended that:

1. Institutions with baccalaureate and master's level programs in statistics should offer a one- or two-semester course in sampling theory. While such a course should be heavily steeped in theory, it should also stress the realities of choosing a design from among several alternative ones, considering both cost and precision. In addition, the course should be taught by a person who has had actual experience in designing a reasonably complex survey or sampling plan. If there is not such a person on its faculty, the institution should explore various means of gaining assistance in

recruiting competent professionals. The institution should, at a minimum, include several lectures given by such a person in its sampling theory course.

2. Institutions with baccalaureate and master's programs in statistics should offer a one- or two-semester course in statistical applications, if they do not already do so. Such a course should include education in design of experiments, quality control and reliability methods, and other specialized statistical procedures, stressing their application to actual data.

3. Baccalaureate and master's programs in statistics should require education in missing/incomplete data techniques. These techniques could be covered in a variety of courses, such as design of experiments, time series analysis, and multivariate analysis. While most graduates have had courses in inferential or analytical techniques, they have had little or no experience in applying these techniques to problems that actually occur. It is rare that one is presented with a data base that is complete and without error. Imputation and validation techniques in an applied setting should be stressed in such a course. Further recommendations on methods for achieving this objective are contained in the report "Preparing Statisticians for Careers in Industry".

4. Institutions should require all majors in statistics to have demonstrated capabilities in the areas of computer programming and statistical software packages. A statistician without such capabilities is not prepared to assume most statistical positions in the labor market - either in government or in private industry.

5. All statistical majors (undergraduate or graduate) should be required to take a one-semester course in technical writing. The most frequent complaint that is levelled against US government statisticians is that they cannot write adequately. Years of hard work on very difficult projects can go for naught if the results and implications cannot be conveyed clearly to senior officials with no technical background.

6. All programs in statistics (undergraduate or graduate) should include at least one course in which the student is required to consult with subject-matter specialists. As with technical writing, too often excellent talent is not properly utilized because the

individual has weak communication skills. The statistical major must be able to translate the subject-matter specialist's needs into statistical specifications, conduct the work, and then translate the results and implications back into the specialist's terms.

It should be noted that many of the foregoing suggestions are being implemented in various education programs. By the same token, it is recognized that not every university department could or should offer programs designed specifically for prospective government statisticians.

To suggest that the responsibility for better preparing the statistician for career opportunities in government statistical service is solely the responsibility of the university clearly would be inappropriate. Improved communication and co-operation between the institutions responsible for training and the prospective employers (in this case, the government agencies) must be achieved. It is hoped that the recommendations in this paper will serve as a useful step in promoting an ongoing exchange between academic and government officials.

The need to introduce the university student to "real-life" problems encountered by government statisticians has been a recurring theme. Methods to accomplish this objective were a focus of the American Statistical Association Conference on Transfer of Methodology between Academic and Government Statisticians held in 1978 (see conference report). Although the scope of the ASA conference was considerably broader than the education of statisticians, the conference's recommendations to government statistical offices are consonant with findings of the Committee on Training of Statisticians for Government and were presented as recommendations to the federal government agencies:

1. Government agencies should foster opportunities for collaboration with academic statisticians through such procedures as the employment of consultants, temporary exchanges (see Savage, 1978) of personnel with academic institutions, and study leaves for senior and junior staff members.

2. Government statistical offices should develop materials to describe vexing problems with which they are faced. Government statistical offices should also develop data packages that simulate practical governmental applications, such as case studies. These materials can be used as training aids in universities.

3. Government statistical offices should develop procedures for increasing access of university researchers to government statistics so that appropriate research can be undertaken.

While the report of the Committee on Training of Statisticians for Government, and those that have been prepared earlier, are intended to communicate needs between and among the several involved institutions, their full value will not be realised without an organised follow-through effort. The Committee on Training of Statisticians for Government endorsed the recommendations of its predecessor Committee on Training of Statisticians for Industry, that is, that "The statistical societies and the universities should create programs to further address these issues," and of the Conference on Transfer of Methodology Between Academic and Government Statisticians, that is, that "national associations should seek to encourage the development of data sources which can be used in the university curricula to give experience in practical statistical education" and that "national associations should develop a compilation of existing exchange opportunities and arrange for distribution throughout the profession." The consonance of findings from the independent assessments of these groups suggests that implementation of recommendations by the academic institutions, the employers, and the associations is appropriate.

References

[1] Bailar, Barbara A. (1978) *Training Statisticians for the Government*, United States Department of Commerce, Bureau of the Census.

[2] Committee on Training of Statisticians for Government (1982) Preparing Statisticians for Careers in the Federal Government: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Government, *The American Statistician*, 36, 69-81.

[3] Committee on Training of Statisticians for Industry (1980) Preparing Statisticians for Careers in Industry: Report of the ASA Section on Statistical Education Committee on Training of Statisticians for Industry, *The American Statistician*, 34, 65-75.

[4] Fellegi, Ivan (1978) Some Thoughts on Training Needed for Government Statistical Work, in: *Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians*, Washington: American Statistical Association, 43-52.

[5] Malinvaud, Edmond (1978) University Teaching for Future Official Statisticians, in: *Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians*, Washington: American Statistical Association, 53-67.

[6] Murthy, M.N. (1978) Statistical Education, Training, Research and Their Utilization, in: *Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians*, Washington: American Statistical Association, 68-93.

BEST COPY AVAILABLE

[7] Savage, I. Richard (1978) Exchanges of Statisticians, in: *Report of the American Statistical Association Conference on the Transfer of Methodology Between Academic and Government Statisticians*, Washington: American Statistical Association, 116-129.

[8] Task Force on Civil Service Reform, Statistical Policy Coordination Committee (1981) *Executive Development for Statisticians Under the Civil Service Reform Act, United States Department of Commerce, Office of Federal Statistical Policy* and Standards.

[9] U.S. Civil Service Commission (1978) *Occupations of Federal White-Collar Workers*.

PROGRAMMES D'ENSEIGNEMENT ET COUTS DE FORMATION

M Benyaklef, INSEA, Rabat

MISE AU POINT D'UN PROGRAMME D'ENSEIGNEMENT

La mise au point de programmes d'enseignement n'est banale que dans le cas où il s'agit d'apprendre à l'étudiant comment utiliser un ensemble de techniques routinières en vue d'exécuter une tâche assez simplement définissable (certains "recyclages" dans le cadre professionnel tombent dans ce cas).

Dès qu'il s'agit d'objectif complexe, comme la formation d'un "Statisticien-planificateur", les choses ne vont pas d'elles-mêmes.

D'abord une tentative de description du "profil" du cadre à former s'impose. Ensuite un corpus de connaissances et de compétences sera dégagé puis une confection sous forme d'unités pédagogiques en résultera. Enfin un système de "Feedback" permettra des réajustements et des adaptations périodiques des programmes. Ce dernier point est fondamental puisqu'aucun profil ni moyen d'y parvenir ne sont immuables.

Objectif de la formation

La première étape est la plus délicate. Il y a lieu de faire une description des tâches que le lauréat du système devra pouvoir exécuter. Cette description sera aussi détaillée que possible. Par example "calculer des indices statistiques" sera précisé selon le cas par exécuter les calculs des indices classiques" ou par "confection et calculs d'indices de prix".

Toujours dans cette même étape, il y a lieu de décrire l'organisation du travail en définissant le rôle de l'individu au sein du groupe : Comment l'apport de l'individu s'intègre aux autres pour produire les résultats recherchés.

La brièveté de ce que nous en disons n'est pas en rapport avec la complexité de cette première étape.

Corpus des connaissances requises

La deuxième étape consiste à passer en revue les outils nécessaires à l'exécution des tâches voulues. Ceci amène de manière récurrente à évoquer des connaissances prérequises à tel ou tel concept, jusqu'à réduction à un seuil minimal. Ce dernier est le "niveau d'entrée". Deux points sont à noter : Primo, l'algorithme qui dégagera les techniques spécifiquement orientées vers les besoins exprimés ne devra pas nous faire oublier la "formation générale" qui donne une dimension humaine et une ouverture indispensables; Secondo, entraîner quelqu'un à apprendre et rechercher par lui-même est un investissement précieux pour garantir que sa formation ne s'arrêtera pas avec l'Ecole.

L'exemple de la formation d'un statisticien est édifiant à ce sujet. En effet, étant appelé à collaborer aux questions les plus diverses, la connaissance statistique seule peut être un handicap pour [illegible]

A titre d'illustration, supposons qu'il s'agiss [illegible]
les modèles économétriques. On dégagera d' [illegible]
outils statistiques et les rouages économiques [illegible]
raménera essentiellement à la connaissance de [illegible]
qui s'y rattachent. L'inférence statistique n [illegible]
de techniques sur la méthode des moindres c [illegible]
étude des espaces vectoriels et du calcul mat [illegible]
Le niveau d'entreé se réduit alors à un baga [illegible]
est clair qu'une autre dimension est souhaita [illegible]
celle de l'appréciation de la portée d'un mod [illegible]
des conclusions face aux hypothèses avancée [illegible]
bution des erreurs, etc...). D'ailleurs la se [illegible]
base sur certaines notions qu'il est indispens [illegible]
reste pas moins que l'habilité à "la fouille in [illegible]
cultiver.

Toujours dans cette deuxième étape, il s'agit [illegible]
ment à partir du corpus ainsi défini : arrête [illegible]
pédagogiques ou de recherche, leur durée et [illegible]
méthodes d'enseignement et d'entraînement. [illegible]
des connaissances se pose : est-ce indispens [illegible]
est-il le plus efficace? Il n'existe pas de ré [illegible]
d'activité; Cependant une certaine forme d'ex [illegible]
l'individu ne soit hautement motivé. Il est e [illegible]
le vrai examen, à savoir la performance dans [illegible]
qui ont le mérite d'être significatifs malgré q [illegible]
contrôle.

Feedback et adaptation du système

La troisième étape consiste à mettre au point [illegible]
de formation. A cette fin, on peut concevoir [illegible]
périodiques à l'intention de chacune des part [illegible]

a) les étudiants,

b) les enseignants,

c) les employeurs,

d) les lauréats du système.

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

La production statistique étant un support fondamental à la planification, la qualité des données et leur manipulation exigent un personnel bien formé.

Les cycles d'études à l'INSEA couvrent trois niveaux de formation correspondants à 3 étapes de l'activité statistique : la collecte et la mise en ordre des données, [illegible]

[illegible]

Premier niveau de formation

[illegible]

[illegible]

[illegible]

[illegible]

Ingénieurs d'Application

L'Ingénieur d'Application joue un rôle fondamental dans l'activité statistique. [illegible]

Revenons aux cours proprement dits : l'introduction d'un cours de Statistique-Mathématique (estimations et tests) est précédée par des notions assez développées en statistique descriptive et en calcul des probabilités. D'autre part l'enseignement des mathématiques (Algèbre linéaire et Analyse classique) a le double but de fournir les outils indispensables à la statistique et de consolider la formation [illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

[illegible]

de son choix (Analyse des données, Théorie et Pratique des Sondages, Econométrie, Planification, Statistique et Economie des Ressources Humaines, etc...).

Un travail de recherche personnelle est demandé à l'étudiant pour s'assurer de son habilité à mettre en oeuvre les connaissances qu'il a acquises.

COÛT PAR ÉTUDIANT

La détermination du coût de formation est intéressante pour au moins deux raisons : toute production a un coût qu'il est important de connaître (la "production de lauréats Statisticiens" ne fait pas exception), et toute dépense est à confronter à d'autres alternatives d'allocation des ressources.

Le calcul du coût de formation d'un Statisticien constitue un exercice compliqué. Il en est de même lorsqu'on veut déterminer la "valeur" de toute personne ayant acquis une certaine qualification professionnelle.

Va-t-on comptabiliser les seuls frais nécessités par l'acquisition de cette qualification ou va-t-on inclure les dépenses encourues avant l'accès de l'individu au Centre de Formation? (dont par exemple le prix de l'acquisition d'un enseignement primaire et secondaire) pourquoi pas remonter même plus en arrière? Le "manque à gagner" qui se dégage du fait que l'individu se consacre aux études plutôt qu'au travail rémunéré n'est-il pas à prendre en considération?

L'autonomie nationale que l'on dérive d'une formation sur place, quel qu'en soit le coût, n'a-t-elle pas une valeur à intégrer dans la fonction d'utilité du lauréat?

Il devient donc clair qu'un calcul du coût de formation a besoin d'être canalisé, il doit être entrepris avec des hypothèses de simplification et doit viser un objectif précis. Nous allons nous restreindre donc à la formulation suivante :

(1) A combien se chiffrent l'édification et le fonctionnement du Centre de Formation?

(2) A quelle charge supplémentaire le budget de l'Etat sera-t-il soumis annuellement lorsque l'on décide d'ériger un Centre de Formation?

(3) Le lauréat du système à mettre en place revient-il à combien dans ces conditions?

Une fois que la réponse à la question (1) est donnée, il suffit de retrancher l'aide extérieure éventuelle et d'ajouter ou retrancher les frais de bourse ou les droits d'inscription (ou recettes diverses) selon le cas pour obtenir la réponse à la 2e question. La réponse à la 3e question n'est pas destinée, selon notre optique, à calculer les droits dont l'étudiant devra s'acquitter pour suivre ses études. Elle ne tient pas compte non plus des frais d'hébergement, de nourriture, etc... Le chiffre obtenu traduit plutôt l'effort financier nécessaire à la production d'un lauréat, étant entendu qu'il ne s'agit là que de la valeur ajoutée par le Centre de Formation.

Voyons comment nous pourrions répondre à cette question précise :

On pourra d'abord comptabiliser le coût de l'étude du projet d'implantation du centre. Soit I_1 l'amortissement annuel de ce coût. Les calculs étant faits sur la base d'un taux et d'une période d'amortissement correspondants aux conditions qui prévalent à l'endroit où le Centre sera implanté.

Soit I_2 le loyer annuel ou l'amortissement des coûts de construction des bâtiments qui abritent le Centre.

Soit I_3 le loyer annuel ou l'amortissement des frais de premier établissement, du mobilier et de l'équipement divers (par exemple location ou achat d'ordinateur, frais d'installation et d'abonnement au téléphone, etc...).

La somme $I_1 + I_2 + I_3$ sera désignée par I. Elle représente la charge constante annuelle de l'investissement initial correspondant à la conception, à la construction et à l'équipement du Centre de Formation. Cette approche suppose qu'il n'y a pas de problème de financement initial, le crédit étant supposé disponible.

Tout projet solidement étudié et documenté et dont l'utilité est manifeste ne pourrait rencontrer de difficultés de financement. En fait le plus souvent, des bâtiments existent déjà pour abriter provisoirement le Centre naissant et le problème se trouve momentanément résolu.

Nous passons aux charges de la composante humaine du Centre de Formation, à savoir les Administrateurs et les Enseignants. L'étude ayant dégagé des besoins précis en hommes, qui peuvent varier d'année en année, soit A_t l'ensemble des salaires distribués au personnel administratif à l'année t. Le coût A_t sera en général réparti uniformément sur l'ensemble des étudiants du Centre à moins qu'il y ait des raisons pour qu'il en soit autrement.

Par exemple, si le Centre reçoit plusieurs catégories d'étudiants, certains d'entre-eux pourraient requérir des services administratifs particuliers. Un ajustement devrait être fait en conséquence.

D'autre part, le personnel pourrait administrer la recherche parallèlement à la formation, il ne serait pas alors correct d'imputer les salaires aux besoins de l'enseignement seul. Un coefficient qui estime la part consacrée à la formation (et la part consacrée à la recherche) devrait être introduit. D'une manière générale pour ne pas avoir à nous répéter lorsque nous passerons aux charges d'enseignants, du matériel, etc..., (c'était déjà le cas pour l'investissement initial I), nous supposerons que si un centre a une double mission de formation et de recherche ou si plus généralement diverses missions lui incombent, alors un coefficient relatant la part de la formation dans l'ensemble des activités du Centre devra être fixé. Ce coefficient sera appliqué à chaque composante des coûts engendrés par l'édification et le fonctionnement du Centre dans sa globalité. Pour ce qui est des recettes éventuelles, leur imputation est claire. En fait chaque fois que la dissociation de la partie - "Formation" peut être faite, l'imputation devient immédiate.

Pour ce qui est de la rémunération des enseignants, l'enseignement s'effectue d'abord par un certain nombre d'heures passées avec les étudiants. Que l'enseignant soit rattaché de manière permanente au Centre, ou qu'il vienne simplement enseigner à titre de "vacataire" ou "consultant", il est possible de déterminer le coût horaire de l'enseignement dispensé étant entendu qu'il est tenu compte des rémunérations éventuelles des services annexes.

Une dernière composante du coût de production est celle de la consommation des matériels (non inclus dans I) et des frais généraux de fonctionnement (à l'exception de ceux afférents aux ressources humaines). Soit F_t la charge annuelle de fonctionnement.

Nous avons adopté l'année comme base du calcul mais il est clair que la méthode reste valable pour déterminer les charges d'un séminaire ou d'une session de recyclage qui ne s'étale que sur quelques mois ou quelques semaines. L'indice t devra être regardé plutôt comme unité de mesure quelconque du temps.

Considérons alors un programme de formation s'étalant sur une période t. Le coût de ce programme est :

$P_t = I + A_t + E_t + F_t$ où I, A_t et F_t sont définis ci-dessus et E_t représente la rémunération des enseignements inscrits dans le programme.

Désignons par N_t le nombre d'étudiants ayant suivi le programme de formation au cours de la période. Le coût par étudiant est donc :

$$C_t = \frac{P_t}{N_t}$$

Si un contrôle des connaissances est institué à la fin de la période de formation et si un nombre n_t seulement ($n_t < N_t$) est déclaré "lauréat", le coût par lauréat est donc :

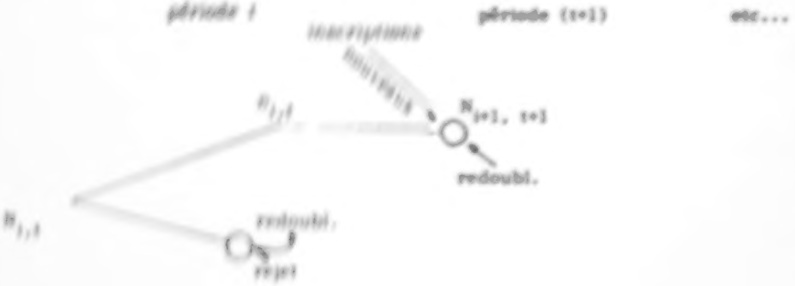
$c_t = \frac{P_t}{n_t}$. Le rapport des coûts $\frac{C_t}{c_t} = \frac{n_t}{N_t}$ exprime le rendement du programme.

FILTRAGE ET RENDEMENT DU SYSTÈME

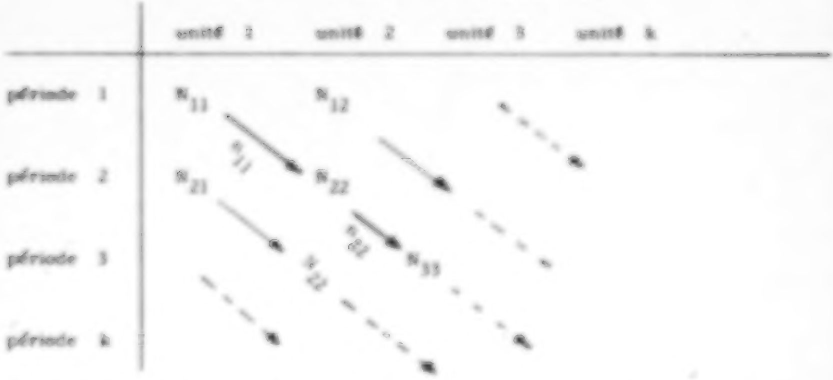
Un programme de formation couvre en général un corpus de connaissances qui se subdivise en unités successives telles que l'inscription à chaque unité requière la réussite à l'unité précédente (exemple des examens trimestriels, semestriels ou annuels). Le contrôle des connaissances ainsi que d'autres facteurs font que le nombre n de lauréats du système est en général inférieur a l'effectif initial N d'étudiants.

Il y a un processus de filtrage à travers le système. A chaque étape, un certain nombre d'étudiants se trouve rejeté du programme prévu par suite d'abandon, d'échec ou de redoublement. Soit N_i l'effectif initial du i^{em} chaînon du système (i^{em} unité) et soit n_i l'effectif final ayant réussi cette unité.

On utilisera un deuxième indice pour désigner la période où ces effectifs sont observés et l'on usera des notations $N_{i,t}$ et $n_{i,t}$. En général on a $N_{i,t} > n_{i,t}$ et $N_{i,t} > n_{i-1,t-1}$. La différence $R_{i,t} = N_{i,t} - n_{i,t}$ est constituée par les abandons, exclusions et redoublements au cours et à la fin de la période t. Quant à la différence $\epsilon_{i,t} = N_{i,t} - n_{i-1,t-1}$ elle est constituée par les redoublements de l'unité i et par les inscriptions éventuelles d'étudiants extérieurs au système et le rejoignant en cours de route. Voici un schéma :



Le schéma des transitions est le suivant



Le rendement du système se dégage à partir des taux de succès. Le vecteur de ces taux étant $(s_{11}, s_{22}, \ldots, s_{kk})$ où $s_{ii} = \frac{n_{ii}}{N_{ii}}$.

Remarquer que les taux utilisés correspondent à une cohorte ou promotion donnée. C'est-à-dire que l'on suit l'évolution d'un groupe d'étudiants inscrits à une année donnée que l'on prend pour origine.

Le rendement peut être appréhendé par le produit $R = s_{11} \times s_{22} \times \ldots s_{kk}$ qui n'est pas très différent du rapport des lauréats au nombre d'inscrits à l'année initiale soit : $r = \frac{n_{kk}}{N_{11}}$ = (r = R si $\varepsilon_{ii} = 0$)

Si l'on veut saisir l'importance des redoublements et des rejets du système, on peut déterminer la durée moyenne $\bar{k}$ nécessaire pour produire un lauréat. La durée normale étant de k, soit r la proportion des étudiants qui terminent leur cycle de formation après k périodes, soit r_1 la proportion de ceux qui terminent après (k+2) périodes etc...., jusqu'à r_m. m étant le nombre maximal de périodes permis en supplément de la durée normale du cycle.

Notons par $r^* = 1 - r - \sum_{\alpha=1}^{m} r_\alpha$ la proportion de ceux qui quittent le système sans jamais obtenir leur diplôme. Nous proposons pour durée moyenne :

$\bar{k} = r.k + \sum_{\alpha=1}^{m} r_\alpha.(k+\alpha) + r^*.2k$ où nous avons choisi arbitrairement le double de la durée du cycle pour tenir compte des abandons et rejets.

On peut également appréhender le rendement de la manière suivante : Soit P_i le coût du programme couvert dans l'unité i, tel qu'il a été calculé plus haut et soit n le nombre de lauréats du système comprenant k unités.
Le coût moyen par lauréat est $\bar{P} = \frac{P_1 + P_2 + \ldots P_k}{n}$. Ce coût moyen par

lauréat pourrait être également saisi en cumulant les coûts moyens par étudiant dans chaque unité

$$\text{Soit } \bar{P} = \frac{P_1}{N_{11}} + \frac{P_2}{N_{22}} + \dots + \frac{P_k}{N_{kk}}$$

Le rendement du système pourrait être traduit par le rapport de ces deux coûts moyens :

$C = \frac{\hat{P}}{\bar{P}}$. Le coefficient $C < 1$. Il est égal à 1 dans le cas où il n'y a pas de déperditions dans le système.

En conclusion, nous pouvons appréhender le rendement du système par l'un des paramètres suivants :

1) $S = s_{11} \cdot s_{22} \cdot \dots \cdot s_{kk}$

2) $\tau = \frac{n_{kk}}{N}$

3) $\bar{k} = \tau . k + \sum_{\alpha=1}^{m} \tau_{\alpha} . (k+\alpha) + \tau^{*} . 2k$

4) $C = \frac{\hat{P}}{\bar{P}}$.

A titre d'exemple, calculons ces paramètres pour le cas du cycle de formation des Ingénieurs d'Application de la Statistique. Ce cycle s'étale sur 3 années ($k = 3$). Pour la cohorte qui a commencé ses études en 1975, les chiffres sont comme suit :

$N_{11} = 92$, $n_{11} = 39$ $S_{11} = 42$ % , $P_1 = 1\,004\,244$ Dh

$N_{22} = 40$, $n_{22} = 32$ $s_{22} = 80$ % , $P_2 = 1\,165\,963$ Dh

$N_{33} = 32$, $n_{33} = 28$ $s_{33} = 88$ % , $P_3 = 566\,057$ Dh

D'autre part $m = 1$ et $\tau_1 = 0,26$. Les calculs donnent :

1) $S = 0,29$
2) $\tau = 0,30$
3) $\bar{k} = (0,30) \times 3 + (0,26) \times 4 + (0,44) \times 6 = 4,58$
4) $C = \frac{\hat{P}}{\bar{P}} = \frac{57\,778}{97\,752} = 0,59$

L'interprétation est que le taux de succès global (pour traverser toutes les étapes du cycle) est de 30 %. La durée moyenne passée par l'étudiant pour sortir lauréat du cycle est près de 4 ans et demi au lieu des 3 années réellement prévues.

Enfin le rapport des coûts moyens entre le système sans déperditions et le système tel qu'il est de 0,59. Dans un certain sens, près de 40 % des dépenses sont à "fonds perdus".

CAPACITÉ IDÉALE ET COÛTS THÉORIQUES

Les moyens aussi bien humains que matériels dont on dispose ne sont pas nécessairement utilisés à l'optimum. Il n'est pas rare de constater dans certaines institutions qu'une salle de classe d'une capacité de 40 n'abrite en fait qu'une dizaine d'étudiants ou bien qu'un spécialiste pouvant encadrer une vingtaine d'auditeurs ne se retrouve en fait qu'avec un faible nombre de disciples. Cette situation résulte parfois de difficultés de recrutement des stagiaires, de l'inadéquation des locaux, de la nature très spécialisée ou avancée de certains enseignements, etc...

La raison peut être tout simplement due à la nature "discontinue" qui lie le nombre d'étudiants aux besoins en salles de classe et en enseignants. En d'autres termes, et à titre d'exemple, on peut utiliser la même salle et le même professeur jusqu'à 20 étudiants mais pour 6 étudiants supplémentaires inscrits dans le même cours, le dédoublement des moyens devient nécessaire, d'où un "Saut" de la fonction coût.

Dans de telles situations, il est naturel de rechercher un taux d'encadrement humain et un taux d'utilisation du matériel qui dans un certain sens atteignent l'optimum. Parmi les contraintes à imposer dans la recherche de cette capacité idéale, il convient de fixer des effectifs qui conduisent à un taux de succès maximal. L'objectif est de garantir l'assimilation des connaissances par les étudiants dans les conditions les meilleures mais sans gaspillage. La résolution de ce programme fait intervenir de nombreux paramètres et dépend largement des spécificités et des besoins de chaque Pays, chaque institution et chaque cycle de formation. La diversité des conditions qui peuvent prévaloir nous recommande de ne pas tenter une résolution dans le cas général.

Si nous nous plaçons à l'optimum, le coût théorique moyen par lauréat sera saisi par un nombre unique à savoir :

$$\bar{P}' = \frac{P'_1 + P'_2 + \dots + P'_h}{n'}$$

où les P'_i sont ceux dont il s'agit plus haut mais avec des effectifs n'_{ik} au lieu de n_{ik} ($n' = n'_{hh}$) qui correspondent à un schéma de transition et à des taux de succès contrôlés.

Le rapport des coûts $\frac{\bar{P}'}{\bar{P}}$ traduit le taux d'efficacité du programme et reflète l'amélioration possible des coûts de formation.

LES ENSEIGNEMENTS STATISTIQUES A L'ENSAE

Ph Tassi et A Trognon, ENSAE, Malakoff Cedex

A - QU'EST QUE L'E.N.S.A.E.?

"L'Ecole Nationale de la Statistique et de l'Administration Economique assure la formation de spécialistes en statistique et en économie appelés à prendre place parmi les cadres de l'Administration, des organismes publics et semi-publics et des entreprises. Elle assure en particulier la formation des fonctionnaires du corps des Administrateurs et du corps des Attachés de l'Institut National de la Statistique et des Etudes Economiques (I.N.S.E.E.).

Elle contribue à la diffusion des connaissances scientifiques et techniques relatives à la méthode statistique, à l'analyse et à la prévision économiques, à l'économétrie, au traitement de l'information, à la démographie et à la sociologie quantitative. Elle concourt à la recherche dans les mêmes domaines".

C'est de cette manière que l'article premier du décret régissant son fonctionnement définit l'E.N.S.A.E.. On y décèle immédiatement la multiplicité des tâches qui lui sont assignées, ce qui se traduit dans les faits par une relative complexité d'organisation que nous allons tenter de résumer.

Premièrement, l'E.N.S.A.E. ne forme pas que des fonctionnaires, la majorité de ses anciens élèves a des responsabilités dans les entreprises françaises ou internationales. Deuxièmement, elle dispense deux niveaux de formation: l'E.N.S.A.E., c'est deux écoles en une. Plus précisément, elle abrite deux divisions qui se distinguent par le recrutement et par le contenu des enseignements.

La Division des Statisticiens Economistes et des Administrateurs (S.E.A.) accueille des élèves qui, ayant reçu une formation scientifique générale d'un niveau élevé sont aptes à assimiler un enseignement fondamental en statistique et en économie et acquérir la pratique des techniques les plus avancées. La division S.E.A. recrute par concours au niveau des classes préparatoires aux Grandes Ecoles (Bac + 2) et, par admission sur titres, d'anciens élèves des Grandes Ecoles scientifiques ou de l'université. Les élèves de cette division passent 2 ou 3 ans à l'Ecole selon leur niveau en mathématique et en économie; ils sont de l'ordre de 250 à l'Ecole, 80 à 90 par promotion.

La Division des Cadres de Gestion Statistique et des Attachés (C.G.S.A.) reçoit des élèves qui, ayant déjà effectué des études supérieures, suivront un enseignement et acquerront la pratique leur permettant de diriger l'exécution de travaux statistiques, de participer à la conception de grandes enquêtes et de collaborer à la préparation de programmes économiques. La division C.G.S.A. recrute par concours au niveau des classes préparatoires et de la deuxième année de Sciences Economiques. Les élèves de cette division effectuent tous deux ans d'études a l'Ecole, ils sont de l'ordre de 160, soit environ 80 par promotion.

D'autre part, depuis 1963, est associée à l'E.N.S.A.E. une école spécialement créée par les communautés européennes afin de former les cadres de la statistique dans les pays francophones en voie de développement : le Centre de formation des Statisticiens Economistes des pays en voie de Développement (C.E.S.D.) Les élèves de ce centre suivent, en plus d'un enseignement spécifique, une part très importante des cours de la division S.E.A.

B - L'ENSEIGNEMENT STATISTIQUE A L'E.N.S.A.E.

Ce cadre rapidement dressé, nous pouvons aborder à présent notre propos principal : l'enseignement de la statistique à l' E.N.S.A.E. Dans cette note rapide, nous n'avons pas dressé un catalogue complet de tous les cours à caractère statistique. Nous avons, au contraire, choisi à titre d'exemple, trois enseignements de la division S.E.A. dont les statuts et les méthodes pédagogiques sont relativement différents : les enseignements de statistique mathématique, d'économétrie et de statistique appliquée.

La statistique mathématique est une discipline fondamentale dont le contenu est assez bien stabilisé.

L'économétrie, par opposition est un domaine en constante évolution dans lequel il faut distinguer un ensemble de connaissances minimales, à acquérir par tous les élèves et les approfondissements s'adressant aux élèves particulièrement intéressés par ce domaine.

Enfin, la statistique appliquée, pour laquelle on ne peut parler ni de niveau minimal à faire connaître, ni de développement constant, procède de la volonté de mettre à l'épreuve de la réalité des données les concepts statistiques et économétriques fondamentaux.

Avant de décrire en détail ces trois enseignements, résumons l'esprit général dans lequel se déroule la scolarité des élèves de la division S.E.A.

En raison des formations antérieures diverses de nos étudiants, la première année dispense un enseignement à dominante mathématique pour ceux des élèves provenant d'une filière à caractère économique et un enseignement principalement économique pour ceux des élèves ayant eu une formation en mathématique.

Tous les élèves de première année acquièrent les connaissances indispensables en calcul et théorie des probabilités. Commencent aussi à un niveau élémentaire et essentiellement pratique des enseignements d'analyse des données et de statistique descriptive.

Année charnière, la deuxième année a pour fonction principale l'enseignement des notions de base de statistique mathématique et d'économie théorique.

Enfin, en troisième année peuvent être mises en application les notions, assimilées lors des années précédentes, soit dans des cours théoriques tel que l'économétrie soit dans des groupes de travail (cf. plus loin) au sein desquels des travaux appliqués sont réalisés, soit encore dans des cours à option.

S'il fallait résumer de manière très synthétique, en quelques mots simples, l'esprit général de chacune des trois années nous pourrions dire :

Première année	:	rattrapage, mise au niveau.
Deuxième année	:	acquisition des connaissances théoriques indispensables.
Troisième année	:	utilisation, mise en oeuvre, application.

LA STATISTIQUE MATHEMATIQUE

Organisation matérielle

Cette matière, enseignée en deuxième année, est constituée par 45 séances(*) de cours magistraux à l'ensemble de la promotion(**) à raison d'une à deux séances par semaine et 34 séances de travaux dirigés (TD). Ces derniers se font par groupe de quinze élèves et ils ont pour vocation de faire résoudre des exercices d'assimilation du cours. Certains TD peuvent nécessiter quelques travaux informatiques ou encore des interprétations de résultats sur listings. D'autres TD, plus volumineux, donnent lieu à notation. Ces TD dits "doubles" sont préparés par deux élèves et présentés oralement aux autres élèves du groupe.

Le contrôle classique des connaissances comprend deux compositions écrites de l'ordre de trois heures, et une interrogation orale.

Contenu de l'enseignement

Destiné à servir de fond à la culture statistique des étudiants et de base aux cours de troisième année où la théorie statistique est un instrument, le cours de statistique mathématique développe les divers éléments de la statistique paramétrique.

(*) Les séances durent 1 heure 10

(**) De l'ordre de 90 élèves

En introduction, cet enseignement définit avec précision la notion de problème statistique en utilisant la théorie de la décision. Il rappelle ou présente ensuite un certain nombre d'outils statistiques généraux (exhaustivité, totalité, information, identification etc...) qui seront utilisés dans le cadre des théories de l'estimation et de tests; enfin la dernière partie de ce cours souligne comment les résultats obtenus dans les parties précédentes s'appliquent au cas du modèle linéaire.

Ce cours bénéficie d'un support écrit cohérent d'un niveau d'abstraction élevé (*). Dans son exposé oral le professeur adopte dans ces conditions un ton moins théorique en mettant l'accent sur des exemples pédagogiques. Il se trouve ainsi allégé des démonstrations mathématiques rigoureuses que l'étudiant peut assimiler par lui-même par référence au cours écrit.

Ainsi dans le cadre de la partie consacrée au modèle linéaire le formalisme adopté le rapproche de celui utilisé en économétrie permettant ainsi une liaison plus facile entre le cours de statistique mathématique de 2ème année et celui d'économétrie de 3ème année.

Une expérience pédagogique passée

Au cours de son histoire, l'enseignement de la statistique mathématique à l'E.N.S.A.E. n'a pas toujours correspondu au [illegible] Ainsi a été réalisée l'éxpérience suivante : e[illegible] cours magistral de théorie des probabilités ([illegible] année) tandis que l'enseignement de la statistique mathématique [illegible] en groupe de travaux dirigés où se mêlaient les exercices et les éléments de cours s'appuyant sur des notes écrites distribuées à l'avance.

Une telle formule présente certains avantages pédagogiques. Elle permet, en particulier, de mieux accorder le niveau du cours à celui des étudiants. Mais, précisément, cette diversité entre les groupes risque d'être néfaste aux enseignements ultérieurs qui s'appuient sur un acquis moyen certain d'une promotion; ce niveau de connaissance peut ne pas être parfaitement défini et très variable d'un groupe d'élèves à un autre. En outre, une telle formule repose fortement sur les notes écrites servant de support aux diverses séances, ce qui entraîne nécessairement une forte rigidité. C'est d'ailleurs essentiellement en raison de la difficulté à la faire évoluer que cette formule a été abandonnée.

(*) Le support de cours est le livre de A. MONFORT : "Cours de Statistique Mathématique", Economica, 1982, France.

LES ENSEIGNEMENTS D'ECONOMETRIE

Organisation

Les enseignements des méthodes statistiques de l'économétrie se déroulent durant l'année terminale, la troisième année. Ils s'articulent autour de deux cours. Le premier, partie intégrante et primordiale du tronc commun, enseigné durant le premier trimestre de l'année scolaire, s'étend sur 22 séances de cours et 10 séances de travaux dirigés. Ce cours dit "Econométrie I" dispense les connaissances fondamentales à l'ensemble de la promotion. Il est sanctionné par un examen écrit.

Le second cours, optionnel, appelé "Econométrie II", est d'un volume de trente séances de cours magistraux et s'adresse aux étudiants les plus motivés par l'économétrie théorique. Ce cours rassemble à peu près le quart d'une promotion. Le contrôle des connaissances repose sur des travaux personnels des étudiants.

Le contenu des enseignements

a) Econométrie I

Ce cours a pour but principal d'apprendre aux étudiants comment les méthodes statistiques qui leur ont été enseignées peuvent être adaptées et précisées pour la spécification et l'estimation des modèles économiques. L'économétrie est une matière en développement continu. Ce cours ne vise donc pas à couvrir l'ensemble des méthodes actuellement disponibles mais plutôt à dispenser une formation de base solide. Il comporte une réflexion critique sur les hypothèses utilisées et les modalités pratiques d'application des techniques étudiées.

Le cours se divise en trois parties. La première est consacrée à une révision approfondie du modèle linéaire et de la méthode des moindres carrés ordinaires, dont les étudiants connaissent déjà bien les principales propriétés théoriques, en insistant sur les problèmes spécifiquement économétriques que soulève leur mise en oeuvre: multicolinéarité, utilisation de variables indicatrices, estimation sous contraintes. Une seconde partie traite du modèle linéaire général et de la méthode des moindres carrés généralisés; notamment dans les cas suivants: hétéroscédasticité, autocorrélation des perturbations, régressions empilées, modèles à double indice. Enfin, dans une troisième partie sont abordées diverses classes de modèles qui ne sont pas réductibles à des modèles linéaires et pour lesquels des méthodes d'estimation particulières ont été mises au point: modèles autorégressifs, modèles dans lesquels la liste des variables explicatives est incomplète, modèle non linéaire, modèles à équations simultanées, modèles à erreurs sur les variables.

Le cours oral est complété par des notes écrites et par des travaux dirigés. Au cours de ceux-ci les étudiants sont amenés à faire passer sur ordinateur les programmes économétriques classiques, à interpréter les résultats d'études économétriques diverses et à approfondir certains points techniques du cours.

Econométrie II

L'Econométrie, nous l'avons déjà souligné, est une matière mouvante sans cesse en renouvellement. La définition du minimum de connaissance que devrait avoir chacun de nos étudiants est délicate. Aussi offrons-nous, aux étudiants qui le désirent une option d'approfondissement portant actuellement sur trois grands thèmes en développement :

- les modèles à équations simultanées
- les modèles pour données individuelles temporelles (panels)
- les modèles pour variables qualitatives.

Chacun de ces trois thèmes est enseigné par un ou des chercheurs de la spécialité. Le niveau élevé de cet enseignement peut être atteint grâce à des supports de cours très détaillés et écrits, sur lesquels peuvent s'appuyer les exposés oraux en ne dégageant que l'esprit principal sans s'attarder sur les techniques de démonstration. Certains de ces supports ne sont pas encore tapés et évoluent au gré de l'avancement de la recherche dans le domaine.

Bien que cet enseignement ne perde jamais de vue les modèles ou les données économiques pour lesquelles les techniques d'estimation et de tests sont élaborées, l'aspect théorique de cet enseignement paraît relativement marqué. Cependant, et dans la mesure du possible, on tente d'associer à chaque procédure nouvelle d'estimation ou de test un programme informatique ou quelques éléments d'algorithmique permettant à chacun d'élaborer son propre programme. L'étudiant dispose dans ces conditions des moyens pour mettre en oeuvre en vraie grandeur, dans le cadre par exemple d'un groupe de travail, les procédures qu'il vient d'apprendre.

Afin de faire manipuler par les étudiants les instruments statistiques utilisés dans ce cours avancé, nous avons tenté cette année une expérience de contrôle des connaissances sur devoirs faits à domicile. Ces devoirs comportaient certaines questions ouvertes où l'étudiant pouvait mettre en valeur des qualités de recherche et d'analyse d'un problème économétrique théorique.

LA STATISTIQUE APPLIQUEE

Structure des enseignements appliqués

Il n'existe pas, dans l'ensemble des enseignements de l'E.N.S.A.E., un et un seul cours d'application de la statistique au sein duquel serait fait le passage obligatoire entre les théories statistiques étudiées et leur mise en oeuvre sur données réelles. La pratique statistique est présente tout au long des trois années d'enseignement: en première année, les élèves réalisent, dans le cadre du contrôle des connaissances du cours d'Analyse des Données, des mémoires en vraie grandeur; en deuxième année, existent, depuis 1977, des groupes dits de "Statistique Appliquée", au sein desquels les étudiants sont confrontés à des problèmes sur données réelles, dont la résolution fait appel aux techniques usuelles de la statistique descriptive; en troisième année, enfin, l'aspect pratique intervient par la présentation d'exemples d'application des méthodes statistiques par des contrôles des connaissances sous forme de mémoire, par la conception et l'exploitation d'une enquête dans le cadre du cours optionnel de Sociologie appliquée, et surtout, par les groupes de travail. Conçus de façon analogue aux groupes de Statistique Appliquée de 2ème année, cette forme pédagogique permet d'aborder des problèmes proches de la recherche, donnant lieu à des travaux théoriques et appliqués où la plupart des cours de l'E.N.S.A.E. peuvent être utilisés.

Nous allons détailler de façon plus approfondie les enseignements de "Statistique Appliquée" et les Groupes de Travail.

Les groupes de "Statistique Appliquée"

La finalité des Groupes de statistique appliquée est, comme cela a été indiqué ci-dessus, de permettre aux étudiants d'aborder des problèmes concrets dont les éléments ne sont pas formalisés.

Cet enseignement commence par huit cours magistraux devant l'ensemble de la promotion, dans lesquels sont abordés les éléments de théorie des indices, de désaisonnalisation et de lissage des séries temporelles, au niveau de la statistique descriptive. Puis les étudiants se répartissent par groupe de 6, chaque groupe choisissant un thème d'étude descriptive parmi plus d'une vingtaine de sujets proposés. Sous la direction d'un animateur connaissant les données et travaillant dans le domaine concerné, les étudiants vont se réunir au rythme de 2 heures tous les quinze jours, pour faire le point de l'avancement de leurs travaux, tout au long de l'année (en moyenne, 10 séances de 2 heures). A la fin de l'année scolaire, un mémoire doit être rédigé, et soutenu oralement.

Les questionnaires d'évaluation de cette forme pédagogique laissent apparaître un intérêt unanime de la part des étudiants, qui découvrent ainsi l'importance des phases antérieures à toute étude formalisée: mise en forme des données, création et structuration de fichiers informatiques, "éclairage" des données par la statistique descriptive.

Il faut également mentionner que la proximité de l'I.N.S.E.E. favorise ce type de travaux, non seulement pour la définition des sujets et la disponibilité des données, mais encore par l'appui de sa structure informatique (terminaux, écrans, logiciels).

A titre d'exemple, nous présentons en annexe 1 la liste des thèmes retenus pour l'année scolaire 1981-1982. Nous l'avons dit, cette forme d'initiation à la description des données est apparue en 1977 dans les enseignements de l'E.N.S.A.E. Jusqu'à cette date, l'apprentissage de la statistique descriptive donnait lieu soit à un cours magistral - voué au désintérêt du public en raison de son manque de support mathématique - soit à des travaux dirigés portant sur les sujets fondamentaux que sont la désaisonnalisation, les indices, la concentration ...

La forme pédagogique des travaux dirigés ne s'est pas révélée adaptée à la matière. La pratique statistique ne s'enseigne pas, mais se découvre et c'est pourquoi la formule des petits groupes autonomes semble la seule possible.

Les groupes de travail

Les groupes de travail constituent, en troisième année, l'analogue des groupes de Statistique Appliquée de 2ème année. L'organisation de l'enseignement est très proche: parmi la quarantaine de thèmes proposés, les étudiants par groupes d'effectif moyen égal à quatre, choisissent un sujet d'études dans des domaines divers: statistique, économie, économétrie, informatique. Chaque groupe se réunit avec le (ou les) animateur(s), deux heures par semaine, et doit en outre, consacrer au moins une demi-journée de travail hebdomadaire à cette activité.

Le travail du groupe est concrétisé par la rédaction d'un mémoire, qui peut parfois donner lieu à publication.

Le fait que le nombre de thèmes proposés soit passé de 6 en 1969, (dans le cadre d'une option), à 43 en 1981, (cette activité étant devenue obligatoire), montre le succès d'un tel enseignement. Ayant dépassé l'analyse descriptive du cours de "Statistique Appliquée", il nécessite la conception et le test de modèles théoriques, et fournit ainsi une première approche d'une direction d'études, qui est souvent l'activité professionnelle d'un ancien élève de la division S.E.A.

La liste des thèmes des groupes de travail proposés en 1981-1982 est fournie en annexe 2.

* * *

L'évolution rapide de la technologie informatique, de son implantation dans les entreprises et, partant, son impact sur la méthodologie statistique, doit être également l'une des préoccupations majeures d'une école formant des statisticiens pour lesquels l'informatique est un outil indispensable.

Dans ce domaine l'E.N.S.A.E. paraît privilégiée par la proximité de l'I.N.S.E.E. L'accessibilité des données, des logiciels (SAS, TSP, BIBLOS) et du matériel (les élèves utilisent un IBM 3033) est une opportunité pour les enseignements à caractère appliqué. Cependant ces facilités ne correspondent pas nécessairement à l'environnement informatique que risqueront de côtoyer nos anciens élèves. Il convient dans l'avenir d'adapter nos enseignements appliqués aux micro-ordinateurs, qui formeront très vraisemblablement l'outil futur du statisticien.

On peut cependant penser que l'apprentissage des logiciels généraux de statistique ne sera pas perdu à terme; l'évolution des tailles de mémoires aidant, certains grands logiciels, par exemple TSP, sont ou seront utilisables sur certains micro-ordinateurs.

La vocation de l'E.N.S.A.E. dépasse largement la déjà très difficile formation des cadres des instituts de statistiques. Les étudiants seront confrontés, dans leur vie professionnelle, à des problèmes très divers, que la plus grande variété des cours à option ne peut évidemment couvrir. Il paraît donc important de fournir à nos élèves un niveau élevé en statistique théorique. Cette culture statistique devrait leur permettre d'être "adaptables" et d'assimiler facilement les connaissances complémentaires dont ils auraient éventuellement besoin par la suite, soit par des lectures personnelles, soit par l'intermédiaire d'actions de formation auxquelles l'E.N.S.A.E. peut participer. Néanmoins, par essence, la statistique est inséparable de l'application. Une école telle que l'E.N.S.A.E. se doit d'être attentive en permanence aux besoins de l'extérieur, savoir quelles sont les méthodes nécessaires aux statisticiens d'entreprises et d'administration; c'est dans cet esprit que se sont développés les stages des étudiants, obligatoires en fin de première année. Sans infléchir véritablement l'enseignement statistique fondamental, de telles informations peuvent orienter le choix des thèmes proposés dans les enseignements de statistique appliquée et les groupes de travail.

Ce double souci de rigueur méthodologique et d'ouverture sur la pratique professionnelle est la préoccupation constante de la direction de l'Ecole et c'est dans cet esprit que celle-ci met actuellement en place une cellule de recherche. Cette cellule, accueillant des jeunes anciens élèves encadrés par des professionnels compétents, devrait être un lieu de rencontre et d'animation intellectuelle qui sera précieux pour l'évolution des enseignements.

ANNEXE 1

THEMES DE STATISTIQUE APPLIQUEE

(Année 1981-1982)

1. Analyse du trafic SNCF (grandes lignes); répartition entre la 1ère classe et la 2ème classe - impact du TGV.

2. Enquête sur les communications écrites et téléphoniques au sein d'une entreprise.

3. Prévisions de ventes.

4. Utilisation du programme CENSUS pour la gestion.

5. La fécondité en France, au niveau départemental, depuis la 2ème guerre mondiale.

6. Analyse d'exportations et d'importations.

7. Modélisation de la diffusion du téléphone en France.

8. Evolution du marché du chauffage.

9. Performances d'un système informatique.

10. La consommation des ménages aux Iles du Cap-Vert.

11. Lien entre activité féminine et fécondité.

12. Caractéristiques des automobilistes en infraction.

13. Influence des traitements de la vigne sur la qualité d'une eau de vie.

14. Une typologie des fraudeurs à l'I.R.P.P.

15. Le chômage régional.

16. Le marché de l'horlogerie.

17. Le rôle de la distribution dans la reconquête du marché intérieur.

18. L'effet de la réduction du temps de travail sur l'emploi dans le secteur du Commerce.

19. Les loisirs en région parisienne.

20. La mobilité pendant le week-end.

21. Prévision de températures journalières.

22. Etude de la validité des tests de personnalité.

23. L'équipement des ménages en biens durables.

ANNEXE 2

GROUPES DE TRAVAIL PROPOSES EN 1981-1982

1. Estimation d'un modèle macroéconomique d'équilibre avec rationnements.

2. La régionalisation: disparités inter-régionales d'épargne et d'investissement.

3. Etude de l'aire d'attraction théorique des espaces verts urbains et péri-urbains.

4. Stratégies de développement et chocs extérieurs.

5. Autonomie du développement économique et ouverture sur l'extérieur dans les pays de l'A.S.E.A.N.

6. Investissement et insertion internationale des économies caractérisées par des salaires réels imparfaitement flexibles.

7. Initiation à l'assembleur I.B.M.

8. Contraintes extérieures au développement: une analyse macroéconomique de déséquilibre.

9. Les méthodes modernes d'analyse des données.

10. Modèles et simulations sur les comptes d'entreprises.

11. La diversification des entreprises industrielles.

12. Analyse des données du projet de comparaisons internationales en Afrique.

13. Pratique de la méthode de Box et Jenkins et développement des programmes.

14. Contrôle monétaire et structures bancaires: évolutions et réformes en cours.

15. Le comportement d'offre.

16. La fécondité différentielle des familles à partir d'un modèle de répartition selon le nombre d'enfants.

17. Rentabilité et stratégie des entreprises françaises.

18. Emploi et productivité: analyse sectorielle.

19. L'analyse statistique et économique de plans d'investissement prévisionnels et de leur validation.

20. Relations entre modèle annuel et modèle trimestriel: application à l'emploi et au cycle prix-salaire.

21. Analyse discriminante.

22. Planification spatiale et marché foncier: appropriation publique des sols.

23. Politique industrielle et préférences nationales des structures.

24. Encombrement du temps et consommation des ménages.

25. Stratégie de développement des pays africains et nouvelle division internationale du travail.

26. Application de la méthode de Box et Jenkins à l'étude prévisionnelle court terme des principaux marchés automobiles européens.

27. Influence du codage et de la distance en analyse des données.

28. Etude de la structure des représentations de l'univers professionnel.

29. Evolution comparée des inégalités sociales en Chine et en Inde.

30. Séries chronologiques de transport de voyageurs.

31. L'endettement international et ses conséquences sur le système financier international.

32. Taux de change et politique économique en France sur longue période (1920-1979).

33. Evaluation des besoins de stationnement auprès des gares de la banlieue parisienne.

34. Spéculation privée et inflation.

35. Application de l'économétrie sur variables qualitatives à une enquête de conjoncture.

36. Dynamiques de long terme et spécialisations des grandes économies occidentales.

37. Incidences régionales des réformes de la politique agricole commune.

38. La conception d'un système d'information.

39. Typologie des entreprises exportatrices dans l'industrie française.

40. Chômage et crise de l'emploi dans les années 1930.

41. Econométrie de la dynamique des revenus et de la consommation.

42. L'équipement des ménages en biens durables.

ISBN 0 946554 00 5 (set of two volumes)

ISBN 0 946554 01 3 (Volume I)

ISBN 0 946554 02 1 (Volume II)

Designed and printed by the University of Sheffield Printing Unit